

According to the Pre Publication Rules, every patent application received by the United States Patent and Trademark Office after November 29, 2000 will be pre-published at eighteen months from the effective filing date. When the application is published the contents, including the sequences, will become prior art.

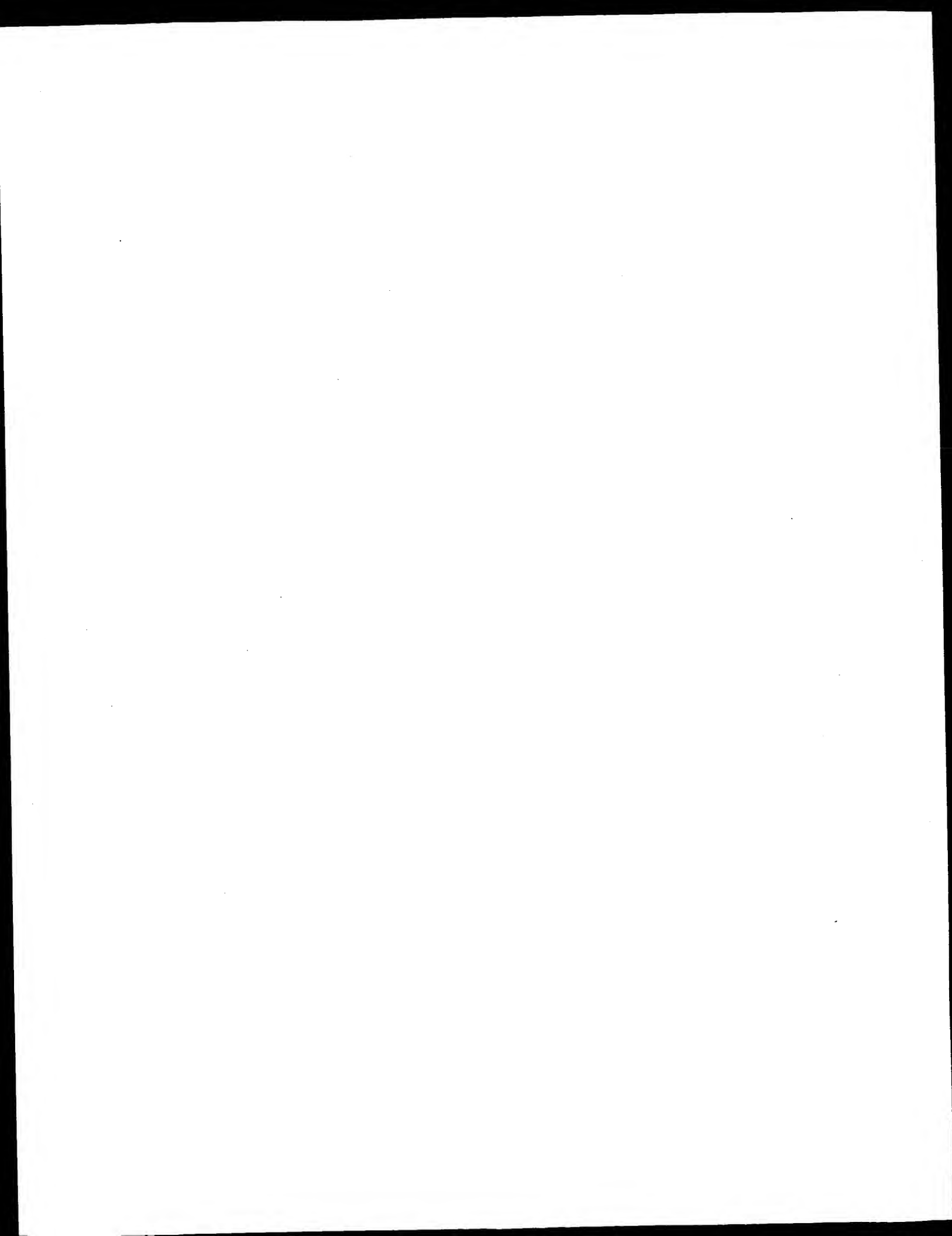
Two new databases have been created to hold the pre-published sequences:

Published_Applications_NA contains nucleic acid sequences; the search results will have the extension **.rnpb**.

Published_Applications_AA contains amino acid sequences; the search results will have the extension **.rapb**.

Each pre-published application is given a unique Publication Number. An example of a Publication Number is US20021234567A1. The "US" indicates the application was a U.S. application. The first 4 digits show the calendar year the application was published. The next 7 digits represent when the application was published. This 7-digit number starts at zero at the beginning of each calendar year. Each application published is given the next number in order. The "A" indicates a utility patent application and the "1" shows that this was the first time the application had been published. If the applicants submit changes to the application, they may request that the changed application be published again. In such instances, the "1" at the end of the number would be replaced by a "2".

Sequences in the PGPub database are public information; it is permissible to leave these results in the case.



GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 28, 2002, 02:51:13 ; Search time 47.9411 Seconds
(without alignments)
7168.683 Million cell updates/sec

Title: US-09-709-103-1
Perfect score: 846
Sequence: 1 atgaactggcgcgatgat.....agcgtgctcatcagctag 846

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 363474 seqs, 203117208 residues
Total number of hits satisfying chosen parameters: 726948

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications_NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	405.8	48.0	3082	10	US-09-778-963A-1
2	278	32.9	405	10	US-09-960-352-10273
3	240.2	28.4	11221	10	US-09-778-963A-3
4	173.2	20.5	368	10	US-09-864-761-21643
5	99	11.7	454	10	US-09-960-352-5830
6	97	11.5	452	10	US-09-960-352-4253
7	92.6	10.9	551	10	US-09-765-298A-25
8	92.6	10.9	570	12	US-10-104-484-1
9	92.6	10.9	570	12	US-10-104-484-3
10	91	10.8	570	10	US-09-765-298A-27
11	74.6	8.8	688	9	US-09-764-868-490
12	74.6	8.8	3346	9	US-09-764-868-67
13	66.6	7.9	1119	9	US-10-067-813-1
14	64.2	7.6	2310	10	US-09-476-242-21
15	64.2	7.6	2535	9	US-09-476-242-13
16	64	7.6	1248	10	US-09-860-846-7
17	64	7.6	1248	10	US-09-861-289-7
18	64	7.6	13613	9	US-09-860-846-3
19	64	7.6	13613	10	US-09-861-289-3

20	61.4	7.3	2298	10	US-09-476-242-22	Sequence 22, Appl
21	61.4	7.3	2298	10	US-09-476-242-23	Sequence 23, Appl
22	61.4	7.3	2298	10	US-09-476-242-24	Sequence 24, Appl
23	61.4	7.3	2529	10	US-09-476-242-14	Sequence 14, Appl
24	61	7.2	2322	10	US-09-476-242-18	Sequence 18, Appl
25	61	7.2	2322	10	US-09-476-242-19	Sequence 19, Appl
26	61	7.2	2322	10	US-09-476-242-20	Sequence 20, Appl
27	61	7.2	2541	10	US-09-476-242-9	Sequence 9, Appl
28	61	7.2	2541	10	US-09-476-242-10	Sequence 10, Appl
29	61	7.2	2541	10	US-09-476-242-11	Sequence 11, Appl
30	61	7.2	2541	10	US-09-476-242-12	Sequence 12, Appl
31	59.8	7.1	309	10	US-09-783-590-4328	Sequence 26, Appl
32	59.4	7.0	2352	10	US-09-476-242-26	Sequence 3, Appl
33	59.2	7.0	1085	9	US-10-067-813-3	Sequence 7960, Ap
34	58.2	6.9	1914	10	US-09-815-242-7960	Sequence 358, App
35	57.4	6.8	1594	10	US-09-969-708-358	Sequence 15, Appl
36	57.4	6.8	2523	10	US-09-476-242-15	Sequence 40, Appl
37	56	6.6	612	10	US-09-972-529-3	Sequence 7866, Ap
38	56	6.6	847	10	US-09-822-849A-40	Sequence 1, Appl
39	56	6.6	1164	10	US-09-815-242-7866	Sequence 17, Appl
40	56	6.6	1224	10	US-09-972-529-1	Sequence 3, Appl
41	56	6.6	2517	10	US-09-476-242-17	Sequence 1, Appl
42	56	6.6	3192	10	US-09-788-654A-1	Sequence 3, Appl
43	55.4	6.5	2310	10	US-09-476-242-3	Sequence 4, Appl
44	55.4	6.5	2316	10	US-09-476-242-4	Sequence 8, Appl
45	55.4	6.5	2316	10	US-09-476-242-8	

ALIGNMENTS

RESULT 1

US-09-778-963A-1
; Sequence 1, Application US/09778963A
; Patent No. US20020115172A1
; GENERAL INFORMATION:
; APPLICANT: NEELAM, Beena et al
; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL001112
; CURRENT APPLICATION NUMBER: US/09/778,963A
; CURRENT FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3082
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-778-963A-1

QY	13	GCATCATCAAGAAGATGCGCGGAGCTGAGTATCCCGCAAGAACTGC	72
Db	208	GCATCATGAAGACTTTGCCAGCGGAACAGCTGCGCGCAAAATCA	267
QY	73	TATCGCATGCTCATCTCGGCTCGTCCAAAGTGGCAAGCGCCATCGTGTCCGCTTC	132
Db	268	TACCGCATGCTGTGCTGGTGCTCTCGGTTGGCAAGAGCTCCATGCTGTCTCGCTTC	327
QY	133	CTCAGCGCGGCTTCGAGGAGCGCTTACACGCTTACATCGAGGACTTCCACCGCAAGTTC	192
Db	328	CTCAATGCGCGCTTTGAGGACAGGTACACACCCATCGAGGACTTCCACCGTAAGTA	387
QY	193	TACTTCATCCGCGGAGGTCTACAGCTCGACATCTTCGACAGTCCGCGCAACCCCG	252
Db	388	TACACATCCGCGGAGCATGTACAGCTCGACATCTGTGATACCTTGGCAACCCCG	447
QY	253	TTCCCGGATCGCGGCGCTCTCCATCTCACAGGACGCTTTCATCCCTGGTTCAGT	312

Query Match 48.0%; Score 405.8; DB 10; Length 3082;
Best Local Similarity 71.4%; Pred. No. 1.7e-76;
Matches 593; Conservative 0; Mismatches 202; Indels 36; Gaps 3;

Query Match 10.9%; Score 92.6; DB 12; Length 570;
Best Local Similarity 57.4%; Pred. No. 3.7e-11;
Matches 167; Conservative 0; Mismatches 124; Indels 0

QY	73	TATCGCATGGT	CATCCTCGGCTCGTCCAAAGCTGGCAAGAGCGGCATCGTCCGCGTTC	132
Db	10			
QY	133	CTCACGGCGCCTT	CGAGGAGCGCTACAGCCTACCATCGAGGACTTCCACCGCAAGTTC	192
Db	70	ATCCAGAACCATTT	TGTGGACGAGTAACGACCCCATATAGAGGATTCTTACCGGAACGAG	129
QY	193	TACTCCATCCGCGGCGAGGTCT	TACCAGTCTGCACATCTCTGCACAGCTCCGGCAACCAACCGG	252
Db	130	GTGGTCAATTGATGGGAGAGCTGCCT	TGTGGACATCTCGATACCGCGGCCACGAGGAG	189
QY	253	TTCGCCGCCATGGGCGCCTCTCCATCTCT	CACAGSAGACCTTTTCATCTCGTGTTTCAGT	312
Db	190	TACAGCGCATCGGAACCAATACAT	GGCACCGGGGAGGCTTCTCTGTGTGTTTGGC	249
QY	313	CTGGACAACCGGCACCTCTCT	TCAGGAGGTGCAGCGGCTCAGGCAGCAGATC	363
Db	250	ATCAACAAACCAAGTCTTTT	CAGGACATCCACAGTACACGGGAGCAGATC	300

```

RESULT 9
US-10-104-484-3
; Sequence 3, Application US/10104484
; Patent No. US20020150566A1
; GENERAL INFORMATION:
; APPLICANT: Guan, Kun-Lian
; APPLICANT: Stewart, Scott
; TITLE OF INVENTION: METHOD OF INHIBITING CANCEROUS CELL PROLIFERATION USING RAS MUTAN
; FILE OF INVENTION: GDP-BOUND CONFORMATION
; FILE REFERENCE: UMICH-0010
; CURRENT APPLICATION NUMBER: US/10/104,484
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER: 60/277,959
; PRIOR FILING DATE: 2001-03-23
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 570
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Nucleotide sequence of the wild type Ras mutant, RasN17N69.
US-10-104-484-3

```

Query Match	10.9%	Score 92.6;	DB 12;	Length 570;
Best Local Similarity	57.4%;	Pred. No. 3.7e-11;		
Matches 167;	Conservative 0;	Mismatches 124;	Indels 0;	Gaps 0;
73	TATCCGATGTCATCTCGGCTCGTCCAAAGTGGGCAAGCGCCATCGTGTGGCGCTTC	132		
DB				
10	TATAAGCTGGTGGTGGCGCGCGGTGTGGGCAAGAAATCGCTGACCATCCAGCTG	69		
QY	CTCACCGCGCGCTTCGAGGAGCGCTACACGCTACCATCGAGGATTTCCACGCCAAGTTC	192		
DB				
70	ATCCAGAACCAATTTGTGGAGCAATACGACCCCACTATAGAGGATTCTCCGGAAGCAG	129		
QY	193	TACTCCATTCGCGCGGAGGTCTACCAAGCTCGACATCCTCGACACGTCCGGCAACCCCG	252	
DB				
130	GTGTGTCATGATGGGAGACGTGCGCTGGACATCTCGATACCGCGGCAGGAGGAG	189		
QY	253	TTCCCGCCCATCGGCGCCTTCCATCTCACGAGAGCGTTTTCATCCTGGTCTTCAGT	312	
DB				
190	TACAGCCCATCGGAACCAAGTACATCGCACCGGGAGGGCTTCCCTGTGTGTGTTGCC	249		
QY	313	CTGGACAACCGCACCTTCCTCGAGAGGTGTCAGCGCTCAGGACGAGATC	363	
DB				
250	ATCAACAACACCAAGTCTTTTTCAGGACATCCACCAGTACAGGAGCAGATC	300		

RESULT 10
US-09-765-298A-27
; Sequence 27, Application US/09765298A

```

: Patent No. US20020137017A1
:
: GENERAL INFORMATION:
:
: APPLICANT: ARONHEIM, AMI
:
: TITLE OF INVENTION: METHOD
:
: FILE REFERENCE: 108387.01
:
: CURRENT APPLICATION NUMBER:
:
: CURRENT FILING DATE: 2001
:
: PRIOR APPLICATION NUMBER:
:
: PRIOR FILING DATE: 1998-07
:
: PRIOR APPLICATION NUMBER:
:
: PRIOR FILING DATE: 1999-01
:
: NUMBER OF SEQ ID NOS: 31
:
: SOFTWARE: Patentin version
:
: SEQ ID NO 27
:
: LENGTH: 570
:
: TYPE: DNA
:
: ORGANISM: Homo sapiens
:
: US-0765-298A-27

```

Query Match	10.8%	Score 91	DB 10	Length 570	
Best Local Similarity	57.0%	Pred. No. 7.9e-11			
Matches 166	Conservative	0	Mismatches 125	Indels	Gaps 0
QY	73	TATCGCATGTCATCTCGGCTGCTCCAAGTGGGCAAGACGGCCATCGTGTGCGCTTC	132		
DB	10	TATAAGCTGGTGGTGGCGCGGGGTGTGGCAAGAGTGCCTGACCATCCAGCTG	69		
QY	133	CTCACCGCGCGCTTCGGAGGAGCCTACACGCCCTACCATCGAGGACTTCAOCGCAAGTTC	192		
DB	70	ATCCAGAACCATTTTGTGGACGAATAGACCCCACTATAGAGGATTCCTACCGGAAGCAG	129		
QY	193	TACTCCATCCGCGCGAGGTCTACCAAGTCCGACATCCTCGACACCTCGCGGCAACCCACCCG	252		
DB	130	GTGGTCAATTGATGGGAGACGTGCCGTGTGGACATCCTGGATACCGCGCGGCTGGAGAG	189		
QY	253	TTCCCGCCCATCGCGGCCCTCTCCATCTCCACAGGAGAGCTTTTCATCCTGGTGTTCAGT	312		
DB	190	TACAGCGCCATGCGGGACCAAGTACATCGCACCGGGGAGGGCTTCCTGTGTGTGTTC	249		
QY	313	CTGGACAAACCGGACTCCTTTCGAGGAGGTGCACGGCTCAGCAGCAGATC	363		
DB	250	ATCAACAACACCAAGTCTTTTGGAGACATCCACCAAGTACAGGAGCAGATC	300		

```

RESULT 11
US-09-764-868-490
; Sequence 490, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT232
; CURRENT APPLICATION NUMBER: US/09/764,868
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 490
; LENGTH: 688
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (579)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (610)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (669)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-868-490

```

Mon Dec 30 09:16:51 2002

us-09-709-103-1-1.rnpb

Query Match 8.8%; Score 74.6; DB 9; Length 688;
 Best Local Similarity 55.1%; Pred. No. 2.2e-07;
 Matches 146; Conservative 0; Mismatches 119; Indels 0; Gaps 0;

QY 102 GTGGGCAAGACGGCCATCGTGTGCGCTTCTCCACGGGCGGCTTCGAGGACGCTACAC 161
 DB 223 GGTGGCAATCTGCCCTTACTGTGCAAGTTTGTCACTGGGACTTTCATTGAGAAATATGA 282

QY 162 GCCTACCATCGAGACTTCACCGCAAGTTTCTACTCCATCCGGGGGAGGTCTTACCAGCT 221
 DB 283 CCCACCATTAAGATTTCTACCCCAAGAGATCGAAGTGGACTCTTCCCGCTCGTGCT 342

QY 222 CGACATCTCTGACAGCTCCGGCAACACCGGTTCCCGGCGCATGGGGGCTCTCCATCCT 281
 DB 343 GGAATTTCTGGACACCGCAGCAAGTGGAGTGTTCCTCCATGAGAGATCTTACATCAA 402

QY 282 CACAGGACGTTTTCATCTCTGTTTCAGTCTGCAACCGGACGCTTCGAGGAGGT 341
 DB 403 AAACGGCCAAAGTTTTCATCTCTGGTTTATAGCTGTTTAAACACAGTCTTTTCAGGATAT 462

QY 342 GCAGGGCTCAGGACGAGATCCTC 366
 DB 463 CRAGCCAATGAGAGATCAAAATGTGC 487

RESULT 12
 US-09-764-868-67
 ; Sequence 67, Application US/09764868
 ; Patent No. US20020168711A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ; FILE REFERENCE: PT232
 ; CURRENT APPLICATION NUMBER: US/09/764,868
 ; Prior application data removed - refer to PALM or file wrapper
 ; NUMBER OF SEQ ID NOS: 1510
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 67
 ; LENGTH: 3346
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: SITE
 ; LOCATION: (2787)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; NAME/KEY: SITE
 ; LOCATION: (2795)
 ; OTHER INFORMATION: n equals a,t,g, or c
 US-09-764-868-67

Query Match 8.8%; Score 74.6; DB 9; Length 3346;
 Best Local Similarity 55.1%; Pred. No. 2.6e-07;
 Matches 146; Conservative 0; Mismatches 119; Indels 0; Gaps 0;

QY 102 GTGGGCAAGACGGCCATCGTGTGCGCTTCTCCACGGGCGGCTTCGAGGACGCTACAC 161
 DB 230 GGTGGCAATCTGCCCTTACTGTGCAAGTTTGTCACTGGGACTTTCATTGAGAAATATGA 289

QY 162 GCCTACCATCGAGACTTCACCGCAAGTTTCTACTCCATCCGGGGGAGGTCTTACCAGCT 221
 DB 290 CCCACCATTAAGATTTCTACCCCAAGAGATCGAAGTGGACTCTTCCCGCTCGTGCT 349

QY 222 CGACATCTCTGACAGCTCCGGCAACACCGGTTCCCGGCGCATGGGGGCTCTCCATCCT 281
 DB 350 GGAATTTCTGGACACCGCAGCAAGTGGAGTGTTCCTCCATGAGAGATCTTACATCAA 409

QY 282 CACAGGACGTTTTCATCTCTGTTTCAGTCTGCAACCGGACGCTTCGAGGAGGT 341
 DB 410 AAACGGCCAAAGTTTTCATCTCTGGTTTATAGCTGTTTAAACACAGTCTTTTCAGGATAT 469

QY 342 GCAGGGCTCAGGACGAGATCCTC 366

DB 470 CRAGCCAATGAGAGATCAAAATGTGC 494

RESULT 13
 US-10-067-813-1
 ; Sequence 1, Application US/10067813
 ; Patent No. US20020156013A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Renauld, Jean-Christophe
 ; APPLICANT: Louahed, Jamila
 ; APPLICANT: Grasso, Luigi
 ; APPLICANT: Levitt, Roy
 ; APPLICANT: Nicolaides, Nicholas
 ; TITLE OF INVENTION: Asthma Associated Factors as Targets for Treating
 ; TITLE OF INVENTION: Atopic Allergies and Related Disorders
 ; FILE REFERENCE: 036870-5071
 ; CURRENT APPLICATION NUMBER: US/10/067,813
 ; CURRENT FILING DATE: 2002-02-08
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/157,247
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1998-09-18
 ; NUMBER OF SEQ ID NOS: 17
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 1
 ; LENGTH: 1119
 ; TYPE: DNA
 ; ORGANISM: Murinae gen. sp.
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (19)...(642)
 US-10-067-813-1

Query Match 7.9%; Score 66.6; DB 9; Length 1119;
 Best Local Similarity 52.7%; Pred. No. 1.1e-05;
 Matches 144; Conservative 0; Mismatches 129; Indels 0; Gaps 0;

QY 103 GTGGGCAAGACGGCCATCGTGTGCGCTTCTCCACGGGCGGCTTCGAGGACGCTACACG 162
 DB 88 GTGGGCAAGAGTGGCTCACTATTTCAGTTTTTCCAGAAAGATCTTTGTGCTGACTACGAC 147

QY 163 CTTACCATCGAGACTTCACCGCAAGTTCTACTCCATCCGGGGGAGGTCTACCAGCTC 222
 DB 148 CCCACCATTAAGACTCTTACCTACCTGAAGCATACAGAGATTGACATCATGTGGCCATCTTG 207

QY 223 GACATCTCTGACAGCTCCGGCAACACCGGTTTCCCGGCGCATCGGCGCTCTCCATCCTC 282
 DB 208 GATGTTCTGGACACACCGCGGCGAGGAGTTCAGTGCCATCGGGAACAATACATGCGC 267

QY 283 ACAGGAGACGTTTTCATCTCTGTTTTCAGTCTGGACAAACCGGACTCTCTCGAGGAGGTG 342
 DB 268 ACAGGGGATGGCTTCTCTCTCTTCTACTCCGTACCGCAAGGCCAGCTTTCGAGCACGTG 327

QY 343 CAGGGGCTCAGGACGAGATCCTTCGACACCAAG 375
 DB 328 GACCGCTTCCACCGACTCATCTTCGCTGTCAAG 360

RESULT 14
 US-09-476-242-21
 ; Sequence 21, Application US/09476242
 ; Patent No. US20020146683A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Barnett, Susan
 ; APPLICANT: BARNETT, Karin
 ; APPLICANT: MARTIN, Eric
 ; TITLE OF INVENTION: MODIFIED HIV ENV POLYPEPTIDES
 ; FILE REFERENCE: 1605.002
 ; CURRENT APPLICATION NUMBER: US/09/476,242
 ; CURRENT FILING DATE: 1999-12-30
 ; NUMBER OF SEQ ID NOS: 26
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 21
 ; LENGTH: 2310
 ; TYPE: DNA

; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Lys121-Vai200;
; OTHER INFORMATION: Asn425-Lys432
US-09-476-242-21

Query Match 7.68; Score 64.2; DB 10; Length 2310;
Best Local Similarity 42.88; Pred. No. 3.7e-05;
Matches 324; Conservative 0; Mismatches 433; Indels 0; Gaps 0;
QY 63 CAAGAACTGCTATCGATGTCATCTCGGCTCGTCCAAAGTGGCAAGCGCCATCGT 122
Db 780 CAACACCTGAGCAGATCGTGACCAAGCTGCGAGGCCAGTTCGGCAACAAGACCATCGT 839
QY 123 CTCGGCTTCCTACCGCGCGCTTCGAGGAGCGCTACAGCCCTACCATCGAGACTTCCA 182
Db 840 GTTCAAGCAGAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 899
QY 183 CCGCAAGTCTTACTTCCATCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 942
Db 900 CGAGTCTTCTACTGCAACAGCAGCCAGCTGTTCACAGCAGCTTGGAAACAACACATCGG 959
QY 243 CAACACCGCTTCCCG 302
Db 960 CCCCACACACACCG 302
QY 303 GGTGTTCACTGAGCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1019
Db 1020 CCCCACACACACCG 362
QY 363 CTTTCGACACCAAGTCTTCCCTCAAGCAACAAACAAAGAGAGAGAGAGAGAGAGAG 1079
Db 1080 CGGCTGCTGTCGACCG 422
QY 423 CATCTGCGGCAACAGGAGTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1139
Db 1140 CCCCACACACACCG 482
QY 483 CGAGCTGTTGGGCGAGCG 1199
Db 1200 GAAGATCAGCCCTTGGCG 542
QY 543 CAGCAGCTGACAGAGTGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1259
Db 1260 GAAGCG 602
QY 603 GAGCCAGACCTGACCG 1319
Db 1320 CATGGCG 662
QY 663 GCTCGGCAACAAGAGTGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1379
Db 1380 GCAGCAGAGAGAGTGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 722
QY 723 CGCTTTGGCATCGTGCG 1439
Db 1440 CGTGTGGGCGATCAAGCAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 782
QY 783 CATCCCGGAGAGCG 1499
Db 1500 CCAGCAGCTGCTGGGCGATCTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1536

RESULT 15
US-09-476-242-13
; Sequence 13, Application US/09476242
; Patent No. US20020146683A1
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: HARTOG, Karin
; APPLICANT: MARTIN, Eric
; TITLE OF INVENTION: MODIFIED HIV ENV POLYPEPTIDES
; FILE REFERENCE: 1605-002

; CURRENT APPLICATION NUMBER: US/09/476,242
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 2535
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Asn425-Lys432
US-09-476-242-13

Query Match 7.68; Score 64.2; DB 10; Length 2535;
Best Local Similarity 42.88; Pred. No. 3.8e-05;
Matches 324; Conservative 0; Mismatches 433; Indels 0; Gaps 0;
QY 63 CAAGAACTGCTATCGATGTCATCTCGGCTCGTCCAAAGTGGCAAGCGCCATCGT 122
Db 1005 CAACACCTGAGCAGATCGTGACCAAGCTGCGAGGCCAGTTCGGCAACAAGACCATCGT 1064
QY 123 CTCGGCTTCCTACCGCGCGCTTCGAGGAGCGCTACAGCCCTACCATCGAGACTTCCA 182
Db 1065 GTTCAAGCAGAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1124
QY 183 CCGCAAGTCTTACTTCCATCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 242
Db 1125 CGAGTCTTCTACTGCAACAGCAGCCAGCTGTTCACAGCAGCTTGGAAACAACACATCGG 1184
QY 243 CAACACCGCTTCCCG 302
Db 1185 CCCCACACACACCG 1244
QY 303 GGTGTTCACTGAGCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 362
Db 1245 CCCCACACACACCG 1304
QY 363 CTTTCGACACCAAGTCTTCCCTCAAGCAACAAACAAAGAGAGAGAGAGAGAGAGAG 422
Db 1305 CGGCTGCTGTCGACCG 1364
QY 423 CATCTGCGGCAACAGGAGTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 482
Db 1365 CCCCACACACACCG 1424
QY 483 CGAGCTGTTGGGCGAGCG 542
Db 1425 GAAGATCAGCCCTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1484
QY 543 CAGCAGCTGACAGAGTGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 602
Db 1485 GAAGCG 1544
QY 603 GAGCCAGACCTGACCG 662
Db 1545 CATGGCG 1604
QY 663 GCTCGGCAACAAGAGTGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 722
Db 1605 GCAGCAGAGAGAGTGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1664
QY 723 CGCTTTGGCATCGTGCG 782
Db 1665 CGTGTGGGCGATCAAGCAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1724
QY 783 CATCCCGGAGAGCG 819
Db 1725 CCAGCAGCTGCTGGGCGATCTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1761

Search completed: December 28, 2002, 05:55:40
Job time : 63.9411 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 30, 2002, 16:35:36 ; Search time 357 Seconds
(without alignment)
2094.295 Million cell updates/sec

Title: US-09-709-103-3F1

Perfect score: 1731

Sequence: 1 GIPSGAGAPSPRAQPEQSP.....IREKASGSAQDKRKCVIS 332

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+pn.model -DEV=xlp
-Q/cgn2_1/USPTO_spool/SULLIVAN3F1/runat 30122002.144153 22127/app_query.fasta_1.519
-DB=N_Geneseq 101002 -QFMT=fastp -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pcio -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=SULLIVAN3F1@cgn.1.179 -runat 30122002.144153 22127 -NCPU=6 -ICPU=3
-NO_XUPXY -NO_WMAP -LARGQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-XGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_101002.*

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1731	100.0	1801	21	AAZ36893	cDNA encoding an a
2	1714	99.0	1740	21	AAZ36914	DNA encoding an ac
3	1713	99.0	1841	20	AAZ23022	Human kd312 polype
4	1692	97.7	1776	21	AAA49177	cDNA encoding huma
5	1651	95.4	3986	20	AAZ23023	Human kd312 genomi
6	1449	83.7	846	21	AAZ36892	cDNA encoding an a
7	1435.5	82.9	1689	20	AAZ23024	Rat kd312 polypept
8	1388.5	80.2	3079	20	AAZ23025	Rat kd312 genomic
9	930.5	53.8	3427	24	ABK71563	Human dithp polynu
10	929.5	53.7	2973	24	ABL92087	Human tumour Endot
11	898	51.9	837	21	AAZ36913	DNA encoding a hom
12	898	51.9	2832	24	ABL92076	Human Tumour Endot
13	867	50.1	3020	24	ABL92134	Mouse Tumour Endot
14	717	41.4	951	23	AAZ36914	DNA encoding novel
15	630.5	36.4	1305	23	ABL07789	Drosophila melanog
16	494	28.5	624	24	ABQ28318	Oligonucleotide fo
17	494	28.5	624	24	ABQ28319	Oligonucleotide fo
18	459	26.5	624	24	ABQ28320	Oligonucleotide fo
19	459	26.5	624	24	ABQ28321	Oligonucleotide fo
20	376	21.7	368	22	ABA51382	Human breast cell
21	376	21.7	368	22	ABA69388	Human foetal liver
22	376	21.7	368	22	ABA36323	Probe #14789 for g
23	376	21.7	368	22	AAK17661	Human brain expres
24	376	21.7	368	22	AAK43477	Human bone marrow
25	376	21.7	368	22	AAI24259	Probe #14192 for g
26	376	21.7	368	22	AAI49541	Probe #18227 used
27	376	21.7	368	22	AAI09818	Probe #9809 used t
28	376	21.7	368	22	ABS17600	Human genome-deriv
29	359.5	20.8	4543	23	ABL07788	Drosophila melanog
30	354.5	20.5	1087	22	AAI60838	Human polynucleoti
31	354.5	20.5	1108	22	AAI59052	Human polynucleoti
32	354.5	20.5	4167	24	AD376405	Human intracellular
33	343	19.8	2340	22	AAH33981	Human colon cancer
34	330.5	19.1	960	23	ABL22541	Drosophila melanog
35	329.5	19.0	3061	24	ABQ72648	Human MDTT encodin
36	319.5	18.5	4933	22	AAK81235	Human immune/haema
37	319.5	18.5	4934	22	AAK81234	Human immune/haema
38	317.5	18.3	2238	22	AAH41603	Human Rap1A, RaIGD
39	317.5	18.3	2238	24	ABL50190	Human Rap1A, RaIGD
40	317.5	18.3	2238	24	ABL50222	Human Rap1A, C-Raf
41	313	18.1	2205	24	ABL50199	Human Rap1A, C-Raf
42	313	18.1	2205	24	ABL50231	Human Rap1A, C-Raf
43	313	18.1	2277	24	ABL50202	Human Rap1A, C-Raf
44	313	18.1	2277	24	ABL50234	Human Rap1A, C-Raf
45	312.5	18.1	3346	22	AAZ27032	cDNA encoding nove

ALIGNMENTS

RESULT 1

AAZ36893

ID AAZ36893 standard; cDNA; 1801 BP.

XX

AC AAZ36893;

XX

DT 13-MAR-2000 (first entry)

XX

DE cDNA encoding an activator of G protein signalling (AGS) protein.

XX

KW Activator of G protein signalling; AGS; ras-related G protein;

KW GTP hydrolysis; G protein activity; pheromone response pathway;

KW G protein-coupled signal transduction; G-gamma selectivity;

XX cellular signal transduction; ss.

OS Homo sapiens.

XX

PH Location/Qualifiers

FT protein"

PN WO9958670-A1.
 XX 18-NOV-1999.
 XX 07-MAY-1999; 99WO-US10151.
 XX 08-MAY-1998; 98US-0084842.
 PR 07-OCT-1998; 98US-0103355.
 XX (CADU-) CADUS PHARM CORP.
 PA Cismowski M, Duzic E;
 PI WPI; 2000-072337/06.
 XX P-PSDB; AAY53924.
 DR A new activator of G protein signalling used to treat disorders
 PT characterized by an aberrant AGS protein activity -
 PT Disclosure; Page 146-148; 162pp; English.
 PS The present sequence encodes an activator of G protein signalling (AGS)
 CC protein. The cDNA sequence was isolated from a human liver cDNA
 CC library. The AGS protein exhibits homology to ras-related G proteins,
 CC and contains alterations in conserved amino acids consistent with a
 CC deficiency in GTP hydrolysis activity. AGS stimulates G protein
 CC activity, G protein-coupled signal transduction and the pheromone
 CC response pathway in a receptor-independent manner. The AGS protein
 CC also shows G-gamma selectivity, as measured by growth assays in
 CC yeast expressing various mammalian G-gamma constructs, and
 CC tissue-specific expression, as measured by Northern blot analysis.
 CC The AGS protein can be used to screen for compounds that modulate
 CC cellular signal transduction. The protein is used to treat disorders
 CC characterized by an aberrant AGS protein activity or AGS nucleic acid
 CC expression.
 XX SQ Sequence 1740 BP; 422 A; 546 C; 483 G; 289 T; 0 other;

Alignment Scores:

Pred. No.: 8,87e-124 Length: 1740
 Score: 1714.00 Matches: 329
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 99.02% Indels: 0
 DB: 21 Gaps: 0

US-09-709-103-3F1 (1-332) x AAZ36914 (1-1740)

QY 4 SerGlyAlaGlyAlaProSerProSerProSerProGlnProGlnSerProProAlaAla 23
 Db 2 AGCGAGCGGAGCGCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 61
 QY 24 HisProAlaCysHisProSerProSerProGlnProSerProAlaAlaMetIleLysLysMetCysPro 43
 Db 62 CACCGCGCGTGCACCCAGCGCGCTCAGCGCTCTCTGCGCTTCTCTGCGCGCGCGCGC 121
 QY 44 ProProSerArgProLysCysProMetLysLeuAlaAlaMetIleLysLysMetCysPro 63
 Db 122 CCGCCCTCGCGCGCGCTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 181
 QY 64 SerAspSerGluLeuSerIleProAlaLysAsnCysTyrArgMetValIleLeuGlySer 83
 Db 182 AGCGACTCGGAGCTGAGTATCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTG 241
 QY 84 SerLysValGlyLysThrAlaIleValSerArgPheLeuThrGlyArgPheGluAspAla 103
 Db 242 TCCAGGTGGGCAAGACG 301
 QY 104 TyrThrProThrIleLeuAspPheHisArgLysPheTyrSerIleArgGlyGluValTyr 123
 Db 302 TACACGCTTACCATCGAGACTTCCACCGCAAGTTCTACTCCCGCGCGCGCGCGCTAC 361

QY 124 GlnLeuAspIleLeuAspThrSerGlyAsnHisProPheProAlaMetArgArgLeuSer 143
 Db 362 CAGCTCGACATCTCGACACAGTCCGGCAACACCCGTTCCCGCCATGCGCGCGCTCTCC 421
 QY 144 IleLeuThrGlyAspValPheIleLeuValPheSerLeuAspAsnArgAspSerPheGlu 163
 Db 422 ATCTCTCACAGGAGAGCTTTTTCATCTCTGTTGTTAGTCTGGACAACCGCGACTCTCTCGAG 481
 QY 164 GluValGlnArgLeuArgGlnGlnIleLeuAspThrLysSerCysLeuLysAsnLysThr 183
 Db 482 GAGGTGCGAGCGCTCAGCGACGAGATCTCGACACCAAGTCTTGGCTCAAGAAACAAACC 541
 QY 184 LysGluAsnValAspValProLeuValIleCysGlyAsnLysGlyAspArgAspPheTyr 203
 Db 542 AAGGAGAACGTGGACCTGCTGCTCATCTCGCGCAACAGGGTGACCGGACTTCTTAC 601
 QY 204 ArgGluValAspGlnArgGluGluLeuValGlyAspValGlyAspProGlnArgCysAla 223
 Db 602 CGCGAGGTGGACCGCGAGATCGAGCAGCTGCTGGCGCAGCAGCCCGCGCGCTCGGCC 661
 QY 224 TyrPheGluIleSerAlaLysLysAsnSerSerLeuAspGlnMetPheArgAlaLeuPhe 243
 Db 662 TACTTCGAGATCTCGCGCAAGAGACAGCAGCTGACCAAGATGTTCCGCGCGCTCTTC 721
 QY 244 AlaMetAlaLysLeuProSerGluMetSerProAspLeuHisArgLysValSerValGln 263
 Db 722 GCCATGCGCAAGTCTCGCGCAGCAGATGAGCCAGACTGACCGCGCAAGGTCTCGGTGCGAG 781
 QY 264 TyrCysAspValLeuHisLysLysAlaLeuArgAsnLysLysLeuLeuArgAlaGlySer 283
 Db 782 TACTTCGAGCTGCTGCAAGAGCGCGTGGGAAACAAGAGCTGCTGGCGCGCGCGCAGC 841
 QY 284 GlyGlyGlyGlyGlyAspProGlyAspAlaPheGlyIleValAlaProPheAlaArg 303
 Db 842 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTGGCATCTGGCACCTTCCGCGCGCGCG 901
 QY 304 ProSerValHisSerAspLeuMetTyrIleArgGluLysAlaSerAlaGlySerGlnAla 323
 Db 902 CCCAGCGTACACAGCGACCTCATGTACATCCGCGAAGAGCGCGCGCGCGCGCGCGCG 961
 QY 324 LysAspLysGluArgCysValIleSer 332
 Db 962 AAGGACAAGGAGCGCTGCTCATCAGC 988
 RESULT 3
 AAZ23022
 ID AAZ23022 standard; cDNA; 1841 BP.
 XX AC AAZ23022;
 XX AC AAZ23022;
 DT 17-JAN-2000 (first entry)
 XX DE Human kd312 polypeptide encoding cDNA.
 XX KW kd312 polypeptide; cancer; immunodeficiency disorder; AIDS; stroke;
 XX KW heart attack; head trauma; neurodegenerative disease; human;
 XX KW Parkinson's disease; Alzheimer's disease; ss.
 XX OS Homo sapiens.
 XX PN WO9950288-A2.
 XX PD 07-OCT-1999.
 XX PF 30-MAR-1999; 99WO-US06993.
 XX PR 31-MAR-1998; 98US-0053374.
 XX PA (AMGE-) AMGEN INC.
 XX PI Yen K;

DR MPI: 1999-601322/51.
 DR P-PSDB; AAY42693.
 XX kd312 polypeptides useful for treating diseases and disorders
 PT associated with alterations in cell proliferation and cell death
 XX
 XX Claim 1, Fig 8, 85pp; English.
 CC The invention provides nucleic acid molecules encoding human and rat
 CC kd312 polypeptides. The kd312 polypeptides can be expressed by standard
 CC recombinant methodology. The kd312 sequences, and the antibodies against
 CC the proteins may be used to treat or diagnose the presence or progression
 CC of various diseases, e.g., cancer, immunodeficiency disorders (e.g. AIDS),
 CC stroke, heart attack, head trauma, and neurodegenerative diseases (e.g.
 CC Parkinson's disease and Alzheimer's disease). The present sequence
 CC represents the human kd312 cDNA sequence.
 XX
 XX Sequence 1841 BP; 398 A; 605 C; 530 G; 308 T; 0 other;
 SQ
 Alignment Scores:
 Pred. No.: 1.13e-123 Length: 1841
 Score: 1713.00 Matches: 329
 Percent Similarity: 99.70% Conservative: 0
 Best Local Similarity: 99.70% Mismatches: 1
 Query Match: 98.96% Indels: 0
 Gaps: 0
 DB:
 US-09-709-103-3f1 (1-332) x AA223022 (1-1841)
 QY 3 Prosergiyalaglyalaproserproserarhlaglnprogluginserrproprola 22
 DB 108 CCGAGCGAAGCGCGAGCGCCCAAGCCGCGCGCCGAGCGCGAGAGCCCTCCAGCC 167
 QY 23 AlaHisproAlaCyHisProSerAspProGlnProLeuSerAlaLeuSerAlaPro 42
 DB 168 GGTCAACCCGCGCTGCCACCCAGCCAGCCCTTCCCTTCCCTTCCGCCCCG 227
 QY 43 ArgProProSerArgProLeuCyProMetLysLeuAlaMetLysLysMetCys 62
 DB 228 CGCCGCGCTCCGCGCCCTCTGCCCATGAACTGGCGCGAGATGATCAAGAAATGTGC 287
 QY 63 ProSerAspSerGlnLeuSerLleProAlaLysAsnCySTyrArgMetValLleLeuGly 82
 DB 288 CCGAGCGACTCGAGCTGATGCCGCGCAAGAACTGTCATGCGATGTCATCTCGGC 347
 QY 83 SerSerLysValGlyLysThrAlaLleValSerArgPheLeuThrGlyArgPheGluasp 102
 DB 348 TCGTCCAGGTGGGCAAGCGGCCATCGTGGCCCTTCTCAACGCGCGCTTCGAGGAC 407
 QY 103 AlaTyrThrProThrLleGlnAspPheHisArgLysPheTyrSerLleArgGlyGluVal 122
 DB 408 GCCTCAACGCTTACATCGAGACTTCCACCCCAAGTTCTTACTCATCCGCGCGAGGTC 467
 QY 123 TyrGlnLeuaspLleLeuaspThrSerGlyAsnHisProPheProAlaMetArgThrLeu 142
 DB 468 TACCAGCTTCGACATCTCGACAGCTCCGCGAACACCCCTTCCCGCATGGGCGCTC 527
 QY 143 SerLleLeuThrGlyAspValPheLleLeuValPheSerLeuAspAsnArgAspSerPhe 162
 DB 528 TCATCTCTCAAGAGAGAGCTTTTCATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 587
 QY 163 GluGluValGlnArgLeuArgGlnGlnLleLeuAspThrLysSerCysLeuLysAsnLys 182
 DB 588 GAGGAGGTGAGCGCGCTCGAGCGAGCATCTTCCGACCCAGAGTCTTCCCTCAAGAAACAA 647
 QY 183 ThrLysGlnLeuValAspValProLeuValLleCysGlyAsnLysGlyAspArgAspPhe 202
 DB 648 ACCAAGAGAGAGCGTGCAGTGCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 707
 QY 203 TyrArgGluValAspGlnArgGlnLleGlnGlnLeuValGlyLysAspProGlnArgCys 222
 DB 708 TACCGGAGGTGAGCGAGCGCGAGATCGAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 767

QY 223 AlaTyrPheGlnLleSerAlaLysLysAsnSerSerLleuaspGlnMetPheArgAlaLeu 242
 DB 768 GCCTACTTCGAGATCTCGCGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 827
 QY 243 PheAlaMetAlaLysLeuProSerGlnLysSerProAspPheHisArgLysValSerVal 262
 DB 828 TTGGCCATGGCCCAAGCTTCCAGCGAGATGAGCCCAAGCTTCGACCGCAAGGATCTCGGTG 887
 QY 263 GlnTyrCysAspValLleuHisLysLysAlaLeuArgAsnLysLysLeuValArgAlaGly 282
 DB 888 CAGTATCTGAGCTGCTCGACAGAGAGCGCTGCGGAAACAAGAGCTGTGCGGCGGCG 947
 QY 283 SerGlyGlyGlyGlyGlyAspProGlyAspAlaPheGlyLleValAlaProPheAlaArg 302
 DB 948 AGCGGCGCGCGCGCGCGAGCCGAGCGCTTGGCATCTGTGCGACCTTTCGCGCGC 1007
 QY 303 ArgProSerValHisSerAspPheMetTyrLleArgGlnLysAlaSerAlaGlySerGln 322
 DB 1008 CGGCGCGAGCTACAG 1067
 QY 323 AlaLysAspLysGlnArgCysValLleSer 332
 DB 1068 GCCAAG 1097
 RESULT 4
 AAA49177
 ID AAA49177 standard; cDNA; 1776 BP.
 AC AAA49177;
 XX
 AC 03-NOV-2000 (first entry)
 XX
 DE cDNA encoding human GTPase associated protein-7.
 XX
 KW Guanine nucleotide binding protein; GTP-binding protein; G-protein;
 KW KW GTPase; GTPase associated protein; GTPAP; cell proliferation;
 KW KW autoimmune; inflammatory; immune system disorder; cancer; AIDS;
 KW KW acquired immune deficiency syndrome; asthma; atherosclerosis;
 KW KW arthritis; systemic lupus erythematosus; psoriasis; human; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 180..1025
 FT /*tag= a
 FT /product= GTPAP7
 PA WO200031263-A2.
 XX
 PD 02-JUN-2000.
 XX
 PD 23-NOV-1999; 99WO-US28013.
 XX
 PF 23-NOV-1998; 98US-0109592.
 PR 04-FEB-1999; 98US-0118610.
 PR 06-APR-1999; 99US-0127990.
 XX
 PA (INCY-) INCYTE PHARM INC.
 XX
 PI Hillman JL, Tang YT, Bandman O, Lal P, Yue H, Lu DM, Baughn MR;
 PI Yang J, Azimzai Y;
 DR MPI: 2000-400073/34.
 DR P-PSDB; AAY9655.
 XX
 PT Human GTPase associated proteins, polynucleotides, and antibodies,
 PT useful for diagnosing, preventing and treating various diseases such as
 PT atherosclerosis, cancer, acquired immune deficiency syndrome (AIDS),
 PT asthma, and autoimmune diseases -
 XX
 PS Claim 9; Page 125-126; 144pp; English.
 XX Human cDNA libraries from various tissues were screened for GTPase

CC associated proteins (GTPAP). The present sequence is cDNA encoding
 CC human GTPAP-7. This sequence was derived from a cDNA library of the
 CC brain tumour tissue from the parietal lobe of a female.
 CC This protein is expressed in reproductive, nervous and
 CC gastrointestinal tissue. The GTPAP proteins may be used to define
 CC agonists and antagonists of GTPAP activity and to generate antibodies
 CC to GTPAP. This means the GTPAP proteins may be useful for treatment or
 CC prevention of diseases associated with GTPAP such as cell proliferation
 CC disorders, autoimmune disorders, inflammatory disorders, immune system
 CC disorders, cancer, AIDS, asthma, atherosclerosis, arthritis, systemic
 CC lupus erythematosus and psoriasis.

XX
 SQ Sequence 1776 BP; 430 A; 565 C; 490 G; 291 T; 0 other;

Alignment Scores:

Pred. No.: 4,63e-122 Length: 1776
 Score: 1692.00 Matches: 326
 Percent Similarity: 98.79% Conservative: 0
 Best Local Similarity: 98.79% Mismatches: 4
 Query Match: 97.75% Indels: 0
 DB: 21 Gaps: 0

US-09-709-103-3F1 (1-332) x AAA49177 (1-1776)

QY 3 ProSerGlyAlaGlyAlaProSerProSerArgAlaGlnProGluGlnSerProProala 22
 Db 33 CCGAGCGAACCAGCGAGCCCAAGCCCGAGCGCGCCAGCCGAGCAGAGCCCTCCAGCC 92
 QY 23 AlaHisProAlaCysHisProSerAspProGlnProLeuSerAlaLeuSerAlaPro 42
 Db 93 GCTCACCCCGGTCGCCACCCAGCCCTCAGCCGCTCTCTGCGCTCTCTCTGCGCCCGC 152
 QY 43 ArgProProSerArgProLeuCysProMetLysLeuAlaAlaMetLysLysMetCys 62
 Db 153 GCGCGCCCTCGCGGCCCTCTCCCAATGAACCTGGCGCGATCATCAAGAGATGTC 212
 QY 63 ProSerAspSerGluLeuSerLysProAlaLysAsnCysTyrArgMetValLysGly 82
 Db 213 CCGAGCGACTCGGAGTGTAGTATCCGCGCAAGAACTGCTATCGCATGTCTCTCGGC 272
 QY 83 SerSerLysValGlyLysThrAlaValSerArgPheLeuThrGlyArgPheGluAsp 102
 Db 273 TCGTCCAGGTGGCAAGACGGCCATCGTGTGCGCTTCTCTCACCGCGCTTCGAGGAC 332
 QY 103 AlaTyrThrProThrIleGluAspPheHisArgLysPheTyrSerLysArgGlyGluVal 122
 Db 333 GCCTACACGCTACCATCGAGACTTCCACGCGAAGTTCTACTCATCCGCGCGAGGTC 392
 QY 123 TyrGlnLeuAspLysLeuAspThrSerGlyAsnHisProPheProAlaMetArgArgLeu 142
 Db 393 TACCAGCTCGACATCTCGACAGCTCGCGCAACCCAGCTTCCCGCCATGCGTGCCTC 452
 QY 143 SerIleLeuThrGlyAspValPheIleLeuValPheSerLeuAspAsnArgAspSerPhe 162
 Db 453 TCCATCTCTCACAGGAGCTTTTCTATCTGTGTTCAGTCTGGACAAACCGGACTCTTC 512
 QY 163 GluGluValGlnArgLeuArgGlnGlnIleLeuAspThrLysSerCysLysLysAsnLys 182
 Db 513 GAGGAGGTGCGAGCGCTCAGCAGCAGATCTCGACACCAAGCTTTCCTCAAGAACAA 572
 QY 183 ThrLysGluAsnValAspValProLeuValLysCysGlyAsnLysGlyAspArgAspPhe 202
 Db 573 ACCAAGAGAGACGTGGACGTGCGCTGCTATCTGCGGCAACAAAGGTTGACCGGACTTC 632
 QY 203 TyrArgGluValAspGlnArgGluIleGluGlnLeuValGlyAspAspProGlnArgCys 222
 Db 633 TACCAGGAGTGGACAGCGAGATCGAGCAGCTGTGTGGCGAGACCCCGAGCGCTGC 692
 QY 223 AlaTyrPheGluIleSerAlaLysLysAsnSerSerIleuAspGlnMetPheArgAlaLeu 242
 Db 693 GCCTACTTCGAGATCTCGGCCAAGAGAACAGCAGCTGAGCAGATGTTCCGCGCGCTC 752
 QY 243 PheAlaMetAlaLysLeuProSerGluMetSerProAspLeuHisArgLysValSerVal 262

Db 753 TTCCCAATGCGCAAGCTGCGCCAGCATGAGCCAGACCTGACCCGCAAGGTCTCGGTG 812
 QY 263 GlnTyrCysAspValLeuHisLysLysAlaLeuArgAsnLysLysLeuLeuArgAlaGly 282
 Db 813 CAGTACTTGCAGCTGCTGCACAAAGAGCGCTGCGGAACAAAGAGCTGCTCGCGGCGCGC 872
 QY 283 SerGlyGlyGlyGlyClyAspProGlyAspAlaPheGlyIleValAlaProPheAlaArg 302
 Db 873 AGCGGCGCGCGCGCGCGCGCGCGCGCGCGCTTTGCGCATGCTGCGCACCTTCGCGCGC 932
 QY 303 ArgProSerValHisSerAspLeuMetTyrIleArgGluLysAlaSerAlaGlySerGln 322
 Db 933 CCGCCCGAGCTATACAGCGACCTCATGTACATCCCGAGAGGCCAGCGCGCGCGAGCCAG 992
 QY 323 AlaLysAspLysGluArgCysValLysSer 332
 Db 993 GCCAAGGACAGGAGCGCTGCTGTCATCAGC 1022

RESULT 5

AAZ23023

ID AAZ23023 standard; DNA; 3986 BP.

XX
 AC AAZ23023;

XX
 DT 17-JAN-2000 (first entry)

XX
 DE Human kd312 genomic DNA sequence.

XX
 KW kd312 polypeptide; cancer; immunodeficiency disorder; AIDS; stroke;
 KW heart attack; head trauma; neurodegenerative disease; human;
 KW Parkinson's disease; Alzheimer's disease; ss.

XX
 OS Homo sapiens.

XX
 PN WO9950288-A2.

XX
 PD 07-OCT-1999.

XX
 PF 30-MAR-1999; 99WO-US06993.

XX
 PR 31-MAR-1998; 98US-0053374.

XX
 PA (AMGE-) AMGEN INC.

XX
 PI Yen K;

XX
 DR WPI; 1999-601322/51.
 P-PSDB; AAY42693.

XX
 PT kd312 polypeptides useful for treating diseases and disorders
 XX associated with alterations in cell proliferation and cell death
 PS Claim 1; Fig 5; 85pp; English.

XX
 CC The invention provides nucleic acid molecules encoding human and rat
 CC kd312 polypeptides. The kd312 polypeptides can be expressed by standard
 CC recombinant methodology. The kd312 sequences, and the antibodies against
 CC the proteins may be used to treat or diagnose the presence or progression
 CC of various diseases, e.g. cancer, immunodeficiency disorders (e.g. AIDS),
 CC stroke, heart attack, head trauma, and neurodegenerative diseases (e.g.
 CC Parkinson's disease and Alzheimer's disease). The present sequence
 CC represents the human kd312 genomic DNA sequence.

XX
 SQ Sequence 3986 BP; 856 A; 1226 C; 1172 G; 732 T; 0 other;

Alignment Scores:

Pred. No.: 1.84e-118 Length: 3986
 Score: 1651.00 Matches: 329
 Percent Similarity: 82.04% Conservative: 0
 Best Local Similarity: 82.04% Mismatches: 1
 Query Match: 95.38% Indels: 71
 DB: 20 Gaps: 1

SQ Sequence 846 BP; 176 A; 288 C; 252 G; 130 T; 0 other;

Alignment Scores:

Pred. No.: 1.35e-103 Length: 846
 Score: 1449.00 Matches: 281
 Percent Similarity: 100.00% Conservativity: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 83.71% Indels: 0
 DB: 21 Gaps: 0

US-09-709-103-3f1 (1-332) x AA2326892 (1-846)

QY 52 MetLysLeuAlaAlaMetIleLysMetCysProSerAspSerGluLeuSerIlePro 71
 Db 1 ATGAACCTGGCGCGGATGATCAAGAAAGATGGCCGAGCTCGGAGCTGAGTATCCG 60
 QY 72 AlaLysAsnCysTyrArgMetValIleLeuGlySerSerLysValGlyLysThrAlaIle 91
 Db 61 GCCAAGAACTGCTATCGCATGGTTCATCTCGGCTCTCCAGGTGGCAAGACGGCCATC 120
 QY 92 ValSerArgPheLeuThrGlyArgPheGluAspAlaTyrThrProThrIleGluAspPhe 111
 Db 121 GTGTCGGCTTCTTCACCGCGCTTCAGAGCGCTCACACGCTTACCATCGAGGACTTC 180
 QY 112 HisArgLysPheThrSerIleArgGlyGluValTyrGlnLeuAspIleLeuAspThrSer 131
 Db 181 CACCGCAAGTTCTACTCATCCCGCGGAGGTCTACCATCGATCGATCTCGACACGTCC 240
 QY 132 GlyAsnHisProPheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIle 151
 Db 241 GGCACACCGCTTCGCCCGCATCGCGGCTCTCCATCTCCACAGGAGAGCTTTTCATC 300
 QY 152 LeuValPheSerLeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGlnGln 171
 Db 301 CTGCTGTTCAGTCTGACACCGCGACTCTCTCGAGGAGGTGCGAGCGCTCAGCGACG 360
 QY 172 IleLeuAspThrLysSerCysLeuLysAsnLysThrLysGluAsnValAspValProLeu 191
 Db 361 ATCTCTCGACACCAAGTCTTGGCTCAAGAACAAACACAGGAGAGCTGGAGCTGCCCTG 420
 QY 192 ValIleCysGlyAsnLysGlyAspArgAspPheTyrArgGluValAspGlnArgGluIle 211
 Db 421 GTCATCTCGCGCAACAGAGGTGACCGGACTTCTACCGGAGGTGACACCGCGAGATC 480
 QY 212 GluGlnLeuValGlyAspAspProGlnArgCysAlaTyrPheGluIleSerAlaLysLys 231
 Db 481 GAGCAGCTGTGGCGACGACCCCGAGCTGCGCTTCTGAGATCTCGGCCAAGAG 540
 QY 232 AsnSerSerLeuAspGlnMetPheArgAlaLeuPheAlaMetAlaLysLeuProSerGlu 251
 Db 541 AACACAGCTGTGGACCATGTTCCGCGCTTTCGCCATGCGCAAGCTGCCAGCGAG 600
 QY 252 MetSerProAspLeuHisArgLysValSerValGlnTyrCysAspValLeuHisLysLys 271
 Db 601 ATGAGCCGAGACCTGACCCGCAAGTCTCGGTGAGTACTGCGACGTGTCACCAAGAG 660
 QY 272 AlaLeuArgAsnLysLysLeuLeuArgAlaGlySerGlyGlyGlyGlyAspProGly 291
 Db 661 CGCTGCGGAAACAAGAAGCTGCTGCGGCGCGGCGAGCGCGCGCGCGCGCGCGCG 720
 QY 292 AspAlaPheGlyIleValAlaProPheAlaArgArgProSerValHisSerAspLeuMet 311
 Db 721 CACGCTTTGGCATGTGGACACCTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 780
 QY 312 TyrIleArgGluLysAlaSerAlaGlySerGlnAlaLysAspLysGluArgCysValIle 331
 Db 781 TATATCGCGAGAGGCGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 840
 QY 332 Ser 332
 Db 841 AGC 843

RESULT 7

AAZ232024

ID AAZ232024 standard; cDNA; 1689 BP.

AC AAZ232024;

DT 17-JAN-2000 (first entry)

XX Rat kd312 polypeptide encoding cDNA.

XX kd312 polypeptide; cancer; immunodeficiency disorder; AIDS; stroke;
 KW heart attack; head trauma; neurodegenerative disease; rat;
 KW Parkinson's disease; Alzheimer's disease; ss.

OS Rattus sp.

XX WO9950288-A2.

XX 07-OCT-1999.

XX 30-MAR-1999; 99WO-US06993.

XX 31-MAR-1998; 98US-0053374.

XX (AMGE-) AMGEN INC.

XX Yen K;

XX WPI; 1999-601322/51.

XX P-PSDB; AAY42694.

XX kd312 polypeptides useful for treating diseases and disorders
 associated with alterations in cell proliferation and cell death -

PS Claim 2; Fig 9; 85pp; English.

XX The invention provides nucleic acid molecules encoding human and rat
 kd312 polypeptides. The kd312 polypeptides can be expressed by standard
 recombinant methodology. The kd312 sequences, and the antibodies against
 of various diseases, e.g. cancer, immunodeficiency disorders (e.g. AIDS),
 stroke, heart attack, head trauma, and neurodegenerative diseases (e.g.
 Parkinson's disease and Alzheimer's disease). The present sequence
 represents the rat kd312 cDNA sequence.

SQ Sequence 1689 BP; 417 A; 462 C; 434 G; 376 T; 0 other;

Alignment Scores:

Pred. No.: 3.44e-102 Length: 1689
 Score: 1435.50 Matches: 290
 Percent Similarity: 91.38% Conservativity: 7
 Best Local Similarity: 89.23% Mismatches: 24
 Query Match: 82.93% Indels: 5
 DB: 20 Gaps: 3

US-09-709-103-3f1 (1-332) x AAZ232024 (1-1689)

QY 11 ProSerArgAlaGlnProGluGlnSerProAlaAlaHisProAlaCysHisProSer 30
 Db 1 CCGCTGCTCTACTCTCAGATTCAGGCCAGCTCGCGGTCGCCGAAGCAAACTCTTCC 60
 QY 31 ---AspProGlnProLeuSerAla-----LeuLeuSerAlaProArgProSerArg 47
 Db 61 ACCACTCGGCGGCTCTGCAGGCTCTTCTCTCAGCCACGATCTGCCCTG-GGG 119
 QY 48 ProLeuCysProMetLysLeuAlaAlaMetIleLysLysMetCysProSerAspSerGlu 67
 Db 120 CCGCTGCTCTCAGATTCAGGCCAGCTCGCGGTCGCCGAAGATGTGCCAAGGACTCTCAA 179
 QY 68 LeuSerIleProAlaLysAsnCysTyrArgMetValIleLeuGlySerSerLysValGly 87
 Db 180 CTGAGTATCCCGCAAGAACTGTCTACAGATGGTCTCTCGGCTCATCCAAAGTGGGC 239
 QY 88 LysThrAlaIleValSerArgPheLeuThrGlyArgPheGluAspAlaTyrThrProThr 107

```

Db 240 AAGACGGCGCATCGTGGCGCTTCTCACGGGCGCTTGAGAGACCCCTTACCC 299
QY 108 ILeGluAAspPheHisArgLysPheTyrSerIleArgGlyIleValTyrGlnLeuAspIle 127
Db 300 ATTGAAGCTTCCACCCGAAAGTTTACTCGATCCGCGGCAAGCTACAGTTGGACAT 359
QY 128 LeuAspThrSerGlyAsnHisProPheProAlaMetArgIleuSerIleLeuThrGly 147
Db 360 CTGGACACATCTGGCAATCATCTCGTTCCCGCATGGGGCGCTCTTATCTCCACAGGA 419
QY 148 AspValPheIleLeuValPheSerLeuAspAsnArgAspSerPheGluGluValGlnArg 167
Db 420 GACGTTTCTATCTGTGTTCAGCTTAGACAACCGCATCTCTTGAGAGAGTCCAAAG 479
QY 168 LeuArgGlnGlnIleLeuAspThrTyrSerCysLeuLysAsnLysThrLysGlnAsnVal 187
Db 480 CTCAAACAGCAGATCTTAGACACCAAGTCTCTTCAGAAACCAACCAAGAGATGTG 539
QY 188 AspValProLeuValIleCysGlyAsnLysGlyAspArgAspPheTyrArgIleValAsp 207
Db 540 GACGTGCGGTGTCATTTGGCGGTAAACAAGGACCGGACCTTCTACCGGAGGTGAG 599
QY 208 GlnArgGlnIleGluGlnLeuValGlyAspAspProGlnArgCysAlaTyrPheGluIle 227
Db 600 CAGCGGAGATGAGCGAGCTGTGGCGCATGACCTTCAGCGTTGTGCTTACCTGAGATC 659
QY 228 SerAlaLysLysAsnSerSerLeuAspGlnMetPheArgAlaLeuPheAlaMetAlaLys 247
Db 660 TCGGCCAAGAAAGATACACCTCGACACAGATCTCCGTCGCTCTTCCATGGCCAG 719
QY 248 LeuProSerGlnMetSerProAspLeuHisArgLysValSerValGlnTyrCysAspVal 267
Db 720 CTGCTTACCGAGATGAGCCCTGACTTGCACCGCAAGTGTCTGTGCAAGTGTGACGTG 779
QY 268 LeuHisLysLysValAlaLeuArgAsnLysLysLeuLeuArgLysGlyGlyGly 287
Db 780 CTGCACAAAGGCTGTGAGAAACAAGCTTCTGCTGCGGGGAGC---GGAGGTGGG 836
QY 288 GlnAspProGlnAspAlaPheGlyTyrLeuAlaProPheAlaArgArgProSerValHis 307
Db 837 GCGGACACAGGAGATGCCCTTGGCATCTTGGCGGCCCTTGTCTCCACAGCTTACGTCAT 896
QY 308 SerAspLeuMetTyrIleArgGlnLysAlaSerIleArgSerGlnAlaLysAspLysGln 327
Db 897 ACGGACCTCATGTACATTCGTGAGAAACCACTGTCAACAGCCAGGCTAAGACAAAGAG 956
QY 328 ArgCysValIleSer 332
Db 957 CGCTGTCTCATCACT 971

```

RESULT 8
AA223025
AA223025 standard; DNA; 3079 BP.

XX AA223025;
XX AC
XX 17-JAN-2000 (first entry)
XX DE
XX Rat kd312 genomic DNA sequence.
XX KM
XX kd312 polypeptide; cancer; immunodeficiency disorder; AIDS; stroke;
XX heart attack; head trauma; neurodegenerative disease; rat;
XX Parkinson's disease; Alzheimer's disease; ss.
XX
XX Rattus sp.
XX OS
XX WO9950288-A2.
XX PN
XX 07-OCT-1999.
XX PD
XX 30-MAR-1999; 99WO-US06993.
XX PF
XX

```

PR 31-MAR-1998; 98US-0053374.
XX (AMGE-) AMGEN INC.
XX PA
XX Ven K;
XX WPI; 1999-601322/51.
XX DR P-PSDB; AA42694.
XX kd312 polypeptides useful for treating diseases and disorders
XX associated with alterations in cell proliferation and cell death
XX
XX Claim 2; Fig 7; 85pp; English.
XX
XX The invention provides nucleic acid molecules encoding human and rat
XX kd312 polypeptides. The kd312 polypeptides can be expressed by standard
XX recombinant methodology. The kd312 sequences, and the antibodies against
XX the proteins may be used to treat or diagnose the presence or progression
XX of various diseases, e.g. cancer, immunodeficiency disorders (e.g. AIDS),
XX stroke, heart attack, head trauma, and neurodegenerative diseases (e.g.
XX Parkinson's disease and Alzheimer's disease). The present sequence
XX represents the rat kd312 genomic DNA sequence.
XX
XX Sequence 3079 BP; 708 A; 843 C; 829 G; 699 T; 0 other;
XX
XX Alignment Scores:
XX Pred. No.: 3,11e-98 Length: 3079
XX Score: 1388.50 Matches: 291
XX Percent Similarity: 79.47% Conservative: 7
XX Best Local Similarity: 77.60% Mismatches: 27
XX Query Match: 80.21% Indels: 52
XX DB: 20 Gaps: 4
XX
XX US-09-709-103-3f1 (1-332) x AA223025 (1-3079)
XX
XX 7 GlnAlaProSerProSerArgAlaGlnProGlnGlnSerProProAlaAlaHisProAla 26
XX 160 GGAAGGCTGAGCCCGCTGCTGCTACTCAAGATTCAGGCCAGCTCGGGGCTCCGAAG 219
XX 27 CysHisProSer---AspProGlnProLeuSerAla-----LeuLeuSerLalaProArg 43
XX 220 CCAAGCTCTTCCACCATCTCCGGGCGCCCTGCGACGCTTACCTTCTTCACAGCGCAT 279
XX 44 ProProSerArgProLeuGlyProMetLysLeuAlaAlaMetIleLysLysMetCysPro 63
XX 280 CTGCCCTG-GGGCCCTCTGCTGCCAATGAACCTGCGCGCATGATCAAGAAAGTGTCCCA 338
XX 64 SerAspSerGlnLeuSerIleProAlaLysAsnCysTyrArgMetValIleLeuGlySer 83
XX 339 ACGGACCTGAACTGAGTATCCCGGCAAGACTGTACAGAGTGTGATCTCTCGGCTCA 398
XX 84 SerLysValGlyLysThrAlaIleValSerArgPheLeuThrGlyArgPheGluAspAla 103
XX 399 TCCAAAGTGGCAAGACCGGCATCTGTGCGCTTCTCAAGGGCCGCTTCCAGAGAGCT 458
XX 104 TyrThrProThrIleGluAspPheHisArgLysPheTyrSerIleArgGlyIleValTyr 123
XX 459 TACACCCCTTACATGAAAGCTTCCACCGAAAGTTTACTGATCCCGGGGAGTGTAC 518
XX GlnLeuAspIleLeuAspThrSerGlyAsnHisProPheProAlaMetArgArgLeuSer 143
XX 519 CAGTTGACACTACTGACACATCTGCAATCATCTCGTTCCCGCAAGCGGCGCTCTCT 578
XX 144 Ile----- 144
XX 579 ATCTCTACAGGTGAGTGGGAGACCGACAGGACCGTGGGAGGAGAAATCTGCGGGAGCGG 638
XX 144 ----- 144
XX 639 ATGGGCGGTGTGTGTGCTTGGGCTGTGCTGTCTCTCGGTCTTGACAGTCTCC 698
XX 145 -----LeuThrGlyAspValPheIleLeuValPheSerLeuAsp 157
XX

```


Db	699	CTCACCTTTCCACTCGTTCCTCTGTGA--GGAGACGTTTTTCATTCCTGGTGTTCCAGCTTAGAC	757
Qy	158	AsnArgAspSerPheGluGluValGlnArgGluArgGlnGlnIleLeuAspThrIysSer	177
Db	758	AACCGCGACTCCTTCGAGGAGGTGCAAAAGGCTCAAAAGAGAGATCTCTAGACCAAGTCC	817
Qy	178	CysLeuLysAsnLysThrIysGluAsnValAspValProLeuValIleCysGlyAsnLys	197
Db	818	TGTTCTCAAGAACAAACCAAGAGAATGTGAGCTGCGCTGTCATTTCCGGTAACAA	877
Qy	198	GlyAspArgAspPheTyrArgGluValAspGlnArgGluIleGlnLeuValGlyAsp	217
Db	878	GGGAGCGGGAATTCTACGCGAAGTGAGGAGCGGAGATTGAGCAGCTGGTGGCGCAT	937
Qy	218	AspProGlnArgCysAlaTyrPheGluIleSerAlaLysLysAsnSerSerIleuAspGln	237
Db	938	GACCTCTCAGCTGTGTGCCTACTTCGAGATCTCGCCCAAGAGAGATAGACGCTGACCA	997
Qy	238	MetPheArgAlaLeuPheAlaMetAlaLysLeuProSerGluMetSerProAspLeuHis	257
Db	998	ATGTTCCGTGCGCTCTTTGCCATGGCCCAAGCTGCCTAGCGAGATGAGCCCTGACTTCAC	1057
Qy	258	ArgLysValSerValGlnTyrCysAspValLeuHisLysLysAlaLeuArgAsnLysLys	277
Db	1058	CGCAAGGTGTCTGTGCAGTACTGTGACGTGCTGCACAAAGGCTCTCAGGAACAAGAG	1117
Qy	278	LeuLeuArgAlaGlySerGlyGlyGlyGlyAspProGlyAspAlaPheGlyIleVal	297
Db	1118	CTTCTGCGTGGCGGACG---GGAGGTGGGGCGACCGAGAGATGCTTTGGCATCTTG	1174
Qy	298	AlaProPheAlaArgArgProSerValHisSerAspLeuMetTyrIleArgGluLysAla	317
Db	1175	GGGCGCTTTGCTCGCAGACTACGTGCTAGCTGATAGCGACCTCATGTATCATTCGTGAGAAAACC	1234
Qy	318	SerAlaGlySerGlnAlaLysAspLysGluArgCysValIleSer	332
Db	1235	AGTGTACAGCCAGGCTAAGCAAGGAGCGCTGTGTATCATCT	1279
RESULT	9		
ABK71563			
ID	ABK71563	standard; cDNA; 3427 BP.	
XX	XX		
AC	ABK71563;		
XX	XX		
DT	30-JUL-2002	(first entry)	
XX	XX		
DE	Human dithp polynucleotide #29.		
KW	Human; dithp; diagnostic and therapeutic polynucleotide; gene; ss; bone		
KW	cell proliferative disorder; cancer; tumour; autoimmune disorder; brain		
KW	inflammatory disorder; viral infection; bacterial infection; seizure;		
KW	fungal infection; parasitic infections; developmental disorder; breast;		
KW	endocrine disorder; metabolic disorder; neurological disorder; cervix;		
KW	gastrointestinal disorder; transport disorder; gene therapy; kidney;		
KW	adrenal gland; bone marrow; lung; ovary; pancreas; prostate; spleen;		
XX	skin, testis; thymus.		
OS	Homo sapiens.		
XX	XX		
PN	WO200220754-A2.		
XX	XX		
PD	14-MAR-2002.		
XX	XX		
PF	29-AUG-2001; 2001WO-US27127.		
XX	XX		
PR	05-SEP-2000; 2000US-229747P.		
PR	05-SEP-2000; 2000US-229748P.		
PR	05-SEP-2000; 2000US-229749P.		
PR	05-SEP-2000; 2000US-229750P.		
PR	05-SEP-2000; 2000US-229751P.		
PR	05-SEP-2000; 2000US-230583P.		
PR	06-SEP-2000; 2000US-230505P.		
PR	06-SEP-2000; 2000US-230514P.		

```

Db 537 GAGGCTGGGGCTGAGACAGACAGCTCCCGAGCTCCGGCGCTTCCAGCAGCT 596
QY 44 -----ProPterArgProLeuCyProMetLysLeuAlaMetIleLys 59
Db 597 CTCTGAGCCGTCAGAGAGCCCGCCCATTCCTCC---AGCCC-CGAGCATGATGAAG 652
QY 60 LysMetCysProSerAspSerGluLeuSerIleProAlaLysAsnCystrArgMetVal 79
Db 653 ACTTGTCACGCGGAGATGACAGCTCAGTGTGCCCCCAAAATCATACCGCATGTG 712
QY 80 IleuGlySerSerLysValGlyLysThrAlaIleValSerArgPheLeuThrGlyArg 99
Db 713 GTGCTGGGGCTCTCGGATGGGAGAGCTTCATCGTCTCGCTTCCCTCATAGGCGCG 772
QY 100 PheGluAspAlaIleThrProThrIleGluAspPheHisArgLysPheTyrSerIleArg 119
Db 773 TTTGAGAGACAGTACACACCCCATCGAGGACCTCCACGTAAGGTAATACACATCCGC 832
QY 120 GlyGluValTyrGluLeuAspIleLeuAspThrSerGlyAsnHisPropheProAlaMet 139
Db 833 GCGGACATGTACAGCTCGACATCTCTGATACCTCTGCAACACACCTTCCCGCATG 892
QY 140 ArgArgLeuSerIleLeuThrGlyAspValPheIleLeuValPheSerLeuAspAsnArg 159
Db 893 CGCAGGCTGTCTCTCTCACAGGAGTGTCTTCATCTCGTGTTCAGCCTGGATTAACCGG 952
QY 160 AspSerPheGluGluValGluArgLeuArgGlnGlnIleLeuAspThrLysSerCysLeu 179
Db 953 GAGTCTTGTATAGGTGCAAGCGCTTTCAGAGCAGATCTTGAGGTCAAGTCTCGCTG 1012
QY 180 LysAsnLysThrLysGluAsnValAspValProLeuValIleCysGlyAsnLysGlyAsp 199
Db 1013 AAGAACAAAGCAAGAGAGCGCGGAGCTCCCATCTGATCTGTGGCAACAAGACAC 1072
QY 200 Arg---AspPheTyrArgGluValAspGlnArgGlnIleGluLeuValGlyAspAsp 218
Db 1073 CACGCGAGCTGTGCGCCAGGTGCCACACAGCAGCGCGAGTGTGTGTGCGGCGAC 1132
QY 219 ProGlnArgCysAlaIleTyrPheGluIleSerAlaLysLysAsnSerLeuAspGlnMet 238
Db 1133 ---GAGAACTCGGCTACTTCTGAGGTGTGCGCCAAAGAAACACCAACGTGGAACGATG 1189
QY 239 PheArgAlaLeuPheAlaMetAlaLysLeuProSerGluMetSerProAspLeuHisArg 258
Db 1190 TTCTACGTGCTCTTCAGATGATGCCAAGCTGCCACAGATGAGCCCCCGCTGCATCGC 1249
QY 259 LysValSerValGlnTyrCysAspValLeuHisLysLys-----AlaLeuArgAsnLys 276
Db 1250 AAGATCTCCGTGACGTACGATGACCCCTTCCACCCAGGCCCTTCTGATGCCCGCGTC 1309
QY 277 LysLeuLeuArgAlaGlySerClyGlyGlyGlyAspProGlyAspAlaPheGlyIle 296
Db 1310 AAGGAGATG-----GACGCTTATGCGCATG 1333
QY 297 ValAlaProPheAlaArgArgProSerValHisSerAspLeuMetTyrIleArgGlyLys 316
Db 1334 GTCTCCCTTGGCGCGCGCCGCGCGCTCAACAGTGACTTCAAGTACATCAAGGCCAAG 1393
QY 317 AlaSerAlaGlySerGlnAlaLysAspLysGluArgCysValIle 331
Db 1394 GTCTCTGGAGAGCGCAGGCGCGTGAAGAGGACAAAGTGCACATC 1438

```

```

RESULT 10
ABU92087
ID ABU92087 standard, cDNA; 2973 BP.
XX
XX ABU92087;
AC
XX 30-MAY-2002 (first entry)
DT
XX Human Tumour Endothelial Marker polynucleotide SEQ ID NO 197.
DE
XX Human, mouse; rat; TEM; tumour endothelial marker; NEM; PEM; cytosolatic;
KW

```

```

KW normal endothelial marker; pan-endothelial marker; immunostimulant;
KW antiangiogenic; tumour; neovascularization; vascularised tumour;
KW polycystic kidney disease; diabetes; retinopathy; rheumatoid arthritis;
KW psoriasis; gene; ss.
OS Homo sapiens.
PN WO200210217-A2.
PD 07-FEB-2002.
PF 01-AUG-2001; 2001MO-US24031.
PR 02-AUG-2000; 2000US-222599P.
PR 11-AUG-2000; 2000US-224360P.
PR 11-APR-2001; 2001US-282850P.
PA (UWJO) UNIV JOHNS HOPKINS.
PI St Croix B, Kinzler KW, Vogelstein B;
XX WPI, 2002-291856/33.
DR P-PSDB; ABB90733.
XX
XX An isolated molecule comprising an antibody variable region which
XX specifically binds to an extracellular domain of a tumor endothelial
XX marker (TEM) protein, useful for inhibiting tumor growth.
XX
XX Claim 65; Page 152-153; 331pp; English.
PS
XX
XX The invention relates to an isolated molecule comprising an antibody
XX variable region which specifically binds to an extracellular domain of a
XX tumor endothelial marker (TEM) protein selected from ABB90732, ABB90740,
XX ABB90749, ABB90750 and ABB90769. The antibodies which bind to TEM
XX proteins have cytostatic, immunostimulant and antiangiogenic activity.
XX They are useful for inhibiting tumor growth, polycystic kidney disease,
XX subjects bearing a vascularised tumour, rheumatoid arthritis and psoriasis. Human, mouse
XX and rat TEM genes and the encoded proteins (ABU92075-ABU92141 and
XX ABB90721-ABB90789) are disclosed, as are marker oligonucleotide
XX sequences: tumour endothelial markers (TEM) ABU9196-ABU92041 and
XX ABU92143-ABU92191; normal endothelial markers (NEM) ABU92042-ABU92074;
XX and pan-endothelial markers (PEM) ABU91903-ABU91995.
XX
XX Sequence 2973 BP; 596 A; 967 C; 787 G; 623 T; 0 other;
SQ

```

```

Alignment Scores:
Pred. No.: 1, 2e-62 Length: 2973
Score: 929.50 Matches: 194
Percent Similarity: 71.68% Conservative: 49
Best Local Similarity: 57.23% Mismatches: 71
Query Match: 53.70% Indels: 25
DB: Gaps: 8

```

```

US-09-709-103-3f1 (1-332) x ABU92087 (1-2973)
QY 1 GlyIleProSerGlyAlaGlyAla---ProSerProSerArgAlaGlnProGlnGlnSer 19
Db 37 GCGGTCCCGAGAGGCGCAGGAGGATCCCGCCAGTGAACCGGAGACCAACAGACT 96
QY 20 ProProAlaAlaHisProAlaCysHisProSerAspProGlnProLeuSerAlaLeu--- 38
Db 97 CTGGAGGCTCGGCGGCTGAGAGCAGAGCGAGCTCCCGCAG---CTCCGCGGCTTCCA 153
QY 39 -----LeuSerAlaProArgProProSerArgProLeuCyProMetLysLeuAla 55
Db 154 GCGAGCTCTCTAGCGCGGCGAGGAGCGCGCGCATTCCTCCAGCC-----CGA 204
QY 56 AlaMetIleLysLysMetCysProSerAspSerGluLeuSerIleProAlaLysAsnCys 75
Db 205 GCCATGATGAAGCTTGTCCAGCGGAACTGCAACGCTCAATGTGCCCGCAAAACTCA 264
QY 76 TyrArgMetValIleLeuGlySerSerLysValGlyLysThrAlaIleValSerArgPhe 95

```

Db 265 TACCGCATGTGTGGTGGTCTCTCGGTGGCAGAGCTCATCTGTCTCGTTC 324
 QY 96 LeuThrGlyArgPheGluAspAlaTyrThrProThrIleGluAspPheHisArgLysPhe 115
 Db 325 CTCAATGCCCGCTTTGAGGACCAAGTACACCCACCATCGAGGACTTCCACCGTAAGGTA 384
 QY 116 TyrSerIleArgGlyGluValTyrGlnLeuAspIleLeuAspThrSerGlyAsnHisPro 135
 Db 385 TACAACATCGCGCGACATGTACCATCGATCGACATCTCTGGATACCTCTGGCAACACCC 444
 QY 136 PheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIleLeuValPheSer 155
 Db 445 TTCCCGCCATCGCGAGGCTTCCATCTCACAGGGATGCTTTCATCTCTGGTGTTCAGC 504
 QY 156 LeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGlnGlnIleLeuAspThr 175
 Db 505 CTGGATTAACCGGGAGTCTCTCGATGAGTCAAGCGCTTCAGAGCAGATCTCTGGAGGTC 564
 QY 176 LysSerCysLeuLysAsnLysThrLysGluAsnValAspValProLeuValIleCysGly 195
 Db 565 AAGTCTGCTGGAAGACCAAGCAAGAGCGCGAGCTGCCATGCTCATCTGTGGC 624
 QY 196 AsnLysGlyAspArg---AspPheTyrArgGluValAspGlnArgGluIleGlnLeu 214
 Db 625 AACAAAGAACGACCGCGAGCTGTGCGCGACGAGTCCACACCCAGGCGGAGCTGTG 684
 QY 215 ValGlyAspAspProGlnArgCysAlaTyrPheGluIleSerAlaLysLysAsnSerSer 234
 Db 685 GTGTGCGGCGAC---GAGAACTGGCTACTTTCAGGTGTGCGCAAGAGAACACCAAC 741
 QY 235 LeuAspGlnMetPheArgAlaLeuPheAlaMetAlaLysLeuProSerGluMetSerPro 254
 Db 742 GTGACGAGATGTTCTACGGTCTCTTCAGATGGCCACGCTGCCACAGATGAGGCC 801
 QY 255 AspLeuHisArgLysValSerValGlnTyrCysAspValLeuHisLysLys---Ala 272
 Db 802 GCCCTGATCGCAAGATCTCCGTGAGTACGCTGAGCGCTTCCACCCCGAGCCCTTCTGC 861
 QY 273 LeuArgAsnLysLysLeuLeuArgAlaGlySerGlyGlyGlyGlyAspProGlyAsp 292
 Db 862 ATGCGCGCTCAAGGAGATG-----GAC 885
 QY 293 AlaPheGlyIleValAlaProPheAlaArgProSerValHisSerAspLeuMetTyr 312
 Db 886 GCCTATGGATGTCCTCGCCCTTCGCGCGCCCGCCACGCTCAACAGTACCTCAAGTAC 945
 QY 313 IleArgGluLysAlaSerAlaGlySerGlnAlaLysAspLysGluArgCysValIle 331
 Db 946 ATCAAGGCCAAGTCTTTCGGAGCGCCAGCGCGCTGAGAGGACAAAGTGCACCATC 1002

RESULT 11

AA236913
 ID AA236913 standard; DNA; 837 BP.

AC AA236913;

XX AA236913;

XX 13-MAR-2000 (first entry)

DE DNA encoding a homologue of activator of G protein signalling AGS1.

XX Activator of G protein signalling; AGS, ras-related G protein;
 KW GTP hydrolysis; G protein activity; pheromone response pathway;
 KW G protein-coupled signal transduction; G-gamma selectivity;
 KW cellular signal transduction; AGS1 homologue; ss.

OS Homo sapiens.

XX Key Location/Qualifiers
 FT CDS 1..837

FT /*tag= a
 FT /product= "AGS1 homologue"

PN WO9958670-A1.
 XX
 PD 18-NOV-1999.
 XX
 PF 07-MAY-1999; 99WO-US10151.
 XX
 PR 08-MAY-1998; 98US-0084842.
 PR 07-OCT-1998; 98US-0103355.
 XX
 PA (CADU-) CADUS PHARM CORP.
 XX
 PI Cismowski M, Duzic E;
 XX
 DR WPI; 2000-072337/06.
 XX
 DR P-PSDB; AAY53923.
 XX
 PT A new activator of G protein signalling used to treat disorders
 PT characterized by an aberrant AGS protein activity -
 XX
 XX Example 15; Page 143-144; 162pp; English.
 CC The present sequence encodes a homologue of a human AGS1 (activator
 CC of G protein signalling (AGS)) protein. The AGS cDNA sequence was
 CC isolated from a human liver cDNA library. The AGS protein exhibits
 CC homology to ras-related G proteins, and contains alterations in
 CC conserved amino acids consistent with a deficiency in GTP hydrolysis
 CC activity. AGS stimulates G protein activity, G protein-coupled signal
 CC transduction and the pheromone response pathway in a receptor-independent
 CC manner. The AGS protein also shows G-gamma selectivity, as measured by
 CC growth assays in yeast expressing various mammalian G-gamma constructs,
 CC and tissue-specific expression, as measured by Northern blot analysis.
 CC The AGS protein can be used to screen for compounds that modulate
 CC cellular signal transduction. The protein is used to treat disorders
 CC characterized by an aberrant AGS protein activity or AGS nucleic acid
 CC expression.
 XX

SQ Sequence 837 BP; 177 A; 272 C; 233 G; 155 T; 0 other;

Alignment Scores:

Pred. No.: 7.35e-61 Length: 837
 Score: 898.00 Matches: 179
 Percent Similarity: 74.09% Conservative: 44
 Best Local Similarity: 59.47% Mismatches: 50
 Query Match: 51.88% Indels: 28
 DB: 21 Gaps: 5

US-09-709-103-3f1 (1-332) x AA236913 (1-837)

QY 34 ProLeuSerAlaLeuLeuSerAlaProArgProProSerArgProLeuCysProMetLys 53
 Db 4 CCTGCTTCTCTCGCTTTGTTGTCAGCCCGA-----33
 QY 54 LeuAlaAlaMetIleLysLysMetCysProSerAspSerGluLeuSerIleProAlaLys 73
 Db 34 -----GCCATGATGAAGACTTTGTCCAGCGGGAATGTCAGCTCAGTGTGCCGCAAA 87
 QY 74 AsnCysTyrArgMetValIleLeuGlySerSerLysValGlyLysThrAlaIleValSer 93
 Db 88 AACTCATACCGCATGGTGTCTGGTGGTCTCTCGGTGGCAAGAGCTCCATCTGTGTCT 147
 QY 94 ArgPheLeuThrGlyArgPheGluAspAlaTyrThrProThrIleGluAspPheHisArg 113
 Db 148 CGCTTCTTCAATATGGCCGCTTTGAGGACAGTACACACCCACCATCCAGACCTTCACCGT 207
 QY 114 LysPheTyrSerIleArgGlyGluValTyrGlnLeuAspIleLeuAspThrSerGlyAsn 133
 Db 208 AAGGTATACACATCCCGCGGCGACATGTACAGCTGACATCTCGATACCTCTGGCAAC 267
 QY 134 HisProPheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIleVal 153
 Db 268 CACCCCTTCCCGCGCATCGCAGGCTGTCCATCTCTCACAGGGGATGTCTTATCTGTGTG 327
 QY 154 PheSerLeuAspAsnArgSerPheGluGluValGlnArgLeuArgGlnIleLeu 173

Db 565 ACCAACGTGGACGAGATGTTCTACGTCTTTCAGCATGGCCCACTCCACGAGATG 624
 Qy 253 SerProAspLeuHisArgLysValSerValGlnTyrCysAspValLeuHisLysLys--- 271
 Db 625 AGCCCGCCCTGATGCAAGATCTCCGTGCGAGTACGGTCCACCCAGGCC 684
 Qy 272 ---AlaLeuArgAsnLysLysLeuLeuArgAlaGlySerGlyGlyGlyGlyAspPro 290
 Db 685 TTCTGATCGCGCGCTCAAGGAGATG----- 711
 Qy 291 GlyAspAlaPheGlyLeuValAlaProPheAlaArgProSerValHisSerAspLeu 310
 Db 712 ---GACGCTATGGCATGCTGCGCCCTTCGCCCGCCCGCCCGCTCAACATGACCTC 768
 Qy 311 MetTyrIleArgGluLysAlaSerAlaGlySerGlnAlaLysAspLysGluArgCysVal 330
 Db 769 AGTACATCAAGGCCCAAGTCTCTGGGAAGGCCCGCCGTCGAGAGGACCAAGTGCACC 828
 Qy 331 Ile 331
 Db 829 ATC 831
 RESULT 13
 ID ABL92134
 XX ABL92134 standard; cDNA; 3020 BP.
 AC ABL92134;
 XX 30-MAY-2002 (first entry)
 DT Mouse Tumour Endothelial Marker polynucleotide SEQ ID NO 292.
 DE Human; mouse; rat; TEM; tumour endothelial marker; NEM; PEM; cytostatic;
 KW normal endothelial marker; pan-endothelial marker; immunostimulant;
 KW antiangiogenic; tumour; neoangiogenesis; vascularised tumour;
 KW polycystic kidney disease; diabetes; retinopathy; rheumatoid arthritis;
 KW psoriasis; gene; ss.
 XX Mus musculus.
 OS
 XX
 PN WC0200210217-A2.
 XX
 PD 07-FEB-2002.
 PF
 PF 01-AUG-2001; 2001WO-US24031.
 PR
 PR 02-AUG-2000; 2000US-222599P.
 PR 11-AUG-2000; 2000US-224360P.
 PR 11-APR-2001; 2001US-282850P.
 XX (UYTO) UNIV JOHNS HOPKINS.
 PA
 XX St Croix B, Kinzler KW, Vogelstein B;
 PI WPI; 2002-291856/33.
 DR P-PSDB; ABB90781.
 XX
 PT An isolated molecule comprising an antibody variable region which
 PT specifically binds to an extracellular domain of a tumor endothelial
 PT marker (TEM) protein, useful for inhibiting tumor growth.
 PS Disclosure; Page 294-295; 331pp; English.
 CC The invention relates to an isolated molecule comprising an antibody
 CC variable region which specifically binds to an extracellular domain of a
 CC tumour endothelial marker (TEM) protein selected from ABB90732, ABB90740,
 CC ABB90749, ABB90750 and ABB90769. The antibodies which bind to TEM
 CC proteins have cytostatic, immunostimulant and antiangiogenic activity.
 CC They are useful for inhibiting tumour growth, neoangiogenesis in
 CC subjects bearing a vascularised tumour polycystic kidney disease,
 CC diabetic retinopathy, rheumatoid arthritis and psoriasis. Human, mouse
 CC and rat TEM genes and the encoded proteins (ABL92075-ABL92141 and

CC ABB90721-ABB90789) are disclosed, as are marker oligonucleotide
 CC sequences: tumour endothelial markers (TEM) ABL91996-ABL92041 and
 CC ABL92143-ABL92191; normal endothelial markers (NEM) ABL92042-ABL92074;
 CC and pan-endothelial markers (PEM) ABL91903-ABL91995.
 XX
 SQ Sequence 3020 BP; 690 A; 848 C; 780 G; 702 T; 0 other;
 Alignment Scores:
 Pred. No.: 8.63e-58 Length: 3020
 Score: 867.00 Matches: 168
 Percent Similarity: 77.62% Conservative: 47
 Best Local Similarity: 60.65% Mismatches: 50
 Query Match: 50.09% Indels: 12
 DB: 24 Gaps: 3
 US-09-709-103-3fl (1-332) x ABL92134 (1-3020)

Qy 56 AlaMetIleLysLysMetCysProSerAspSerGluLeuSerIleProAlaLysAsnCys 75
 Db 373 GCCATGATCAAGACCTTGTCCAGTGGGAACGTGCACACTCAATGTGCTGCTAAGAACTCC 432
 Qy 76 TyrArgMetValIleLeuGlySerSerLysValGlyLysThrAlaIleValSerArgPhe 95
 Db 433 TACCCCATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 492
 Qy 96 LeuThrArgPheGluAspAlaTyrThrProThrIleGluAspPheHisArgLysPhe 115
 Db 493 CTCATGGCGCTTTTGGAGACCACTACACCCCTATCGAGGACTTTCATCGCAAGGTG 552
 Qy 116 TyrSerIleArgGlyGluValTyrGlnLeuAspIleLeuAspThrSerGlyAsnHisPro 135
 Db 553 TACAACTATCCACGGGGACATGTACCAAGTGGATATCTTGACACCTCTGGCAACACCCCA 612
 Qy 136 PheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIleLeuValPheSer 155
 Db 613 TTCCTGCCATGCGCGCTCTCCATCTCACAGGAGATGTCTTCATCTGCTGTTTCAGC 672
 Qy 156 LeuAspAsnArgAspSerPheGluValGlnArgLeuArgGlnGlnIleLeuAspThr 175
 Db 673 CTGATAGCGGGAGTCTTTGATGAGGTCAAGCGCTCCAGAACAGATCTCTGGAGGTC 732
 Qy 176 LysSerCysLeuLysAsnLysThrLysGluAsnValAspValProLeuValIleCysGly 195
 Db 733 AAGTCTGCTCGAATATAAACCAAGAGGCGACAGAGTGGCCCATGTGTGATCTGTGGG 792
 Qy 196 AsnLysGlyAspArg---AspPheTyrArgGluValAspGlnArgGluIleGluGlnLeu 214
 Db 793 AACAAAGATGACACAGTGTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCTG 852
 Qy 215 ValGlyAspAspProGlnArgCysAlaTyrPheGluIleSerAlaLysLysAsnSerSer 234
 Db 853 GTGCTGTGGTAT---GAAACTGCGCTATTTTCAGGTGTGTCAGCAAGAAAGAACTAAT 299
 Qy 235 LeuAspGlnMetPheArgAlaLeuPheAlaMetAlaLysLeuProSerGluMetSerPro 254
 Db 910 GTGACAGAGATGTTCTATGTGCTGTTTCAGCATGCGCAAGCTGCCCCCATGATGAGCCCT 969
 Qy 255 AspLeuHisArgLysValSerValGlnTyrCysAspValLeuHisLysLysAlaLeuArg 274
 Db 970 GCACTGCACCATAGATCTCCGTGCGAGTACGGCGATGCTTTTTCACCCCGCGCTTCTGCG 1029
 Qy 275 AsnLysLysLeuLeuArgAlaGlySerGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 1099
 Db 1030 ATGGCTGCCACTAAGGTGCGCAGGT-----GCCTAT 1059
 Qy 295 GlyTleValAlaProPheAlaArgProSerValHisSerAspLeuMetTyrIleArg 314
 Db 1060 GGCACTGCTCACCTTTGCCGAGCGCCCGAGTGTCAACAGTACCTCAAGTACATCAAG 1119
 Qy 315 GluLysAlaSerAlaGlySerGlnAlaLysAspLysGluArgCysValIle 331
 Db 1120 GCCAAGGTCTCTACGGAGGGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1170

RESULT 14

AAS90571 standard; cDNA; 951 BP.

AAS90571;

13-FEB-2002 (first entry)

DNA encoding novel human diagnostic protein #26375.

Human; chromosome mapping; gene mapping; gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder; ss.

Homo sapiens.

MO2001.75067-A2.

11-OCT-2001.

30-MAR-2001; 2001WO-US08631.

31-MAR-2000; 2000US-0540217.

23-AUG-2000; 2000US-0649167.

(HSE-) HSEQ INC.

Drmanac RT, Liu C, Tang YT;

MPI; 2001-639362/73.

P-PSGB; ABG6384.

New isolated polynucleotide and encoded polypeptides, useful in

diagnostics, forensics, gene mapping, identification of mutations

responsible for genetic disorders or other traits and to assess

biodiversity -

Claim 1; SEQ ID No 26375; 103pp; English.

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridization probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for biological disorders involving aberrant protein expression or biological activity in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic coding sequences of the invention.

Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

Sequence 951 BP; 202 A; 291 C; 293 G; 165 T; 0 other;

Alignment Scores:

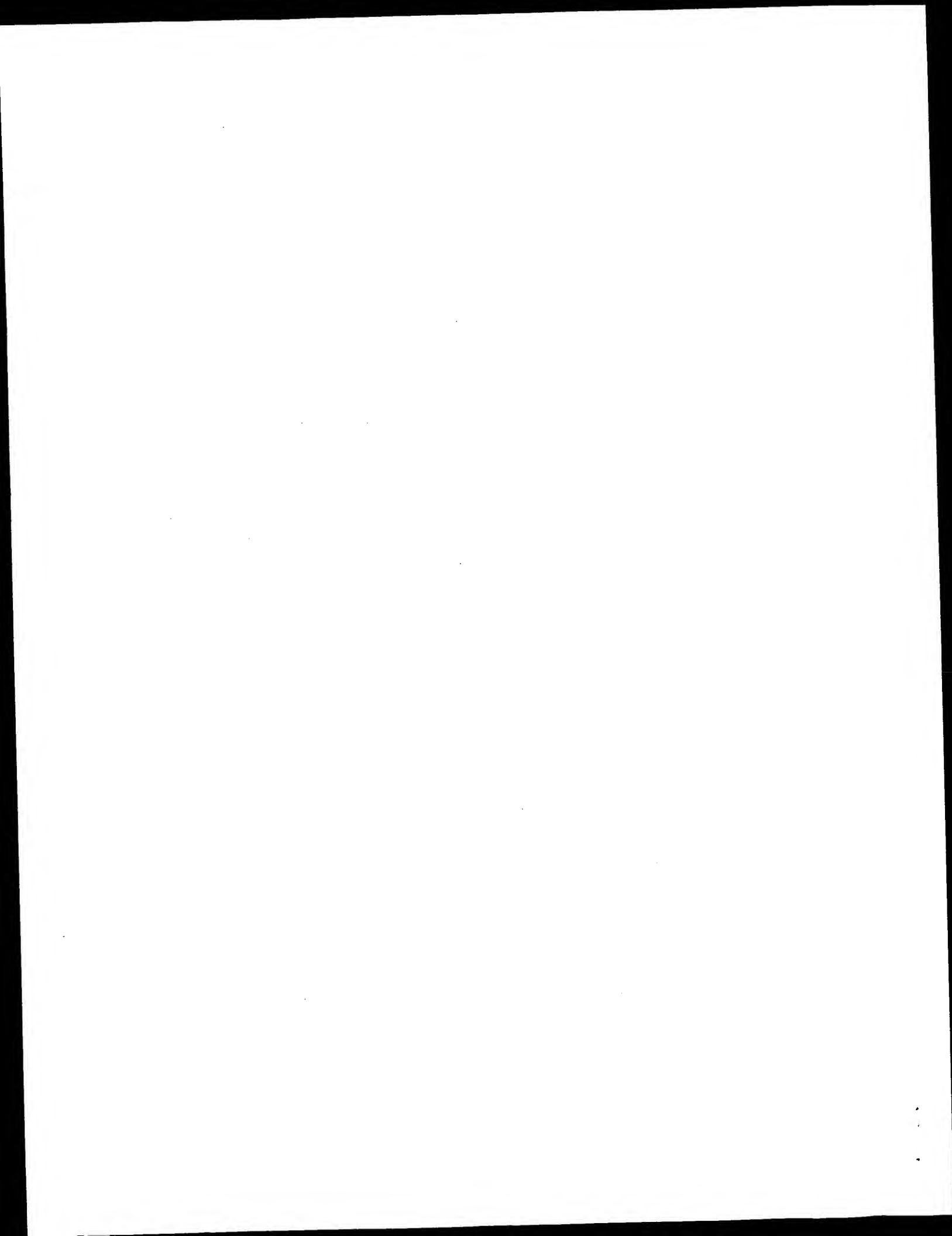
Pred. No.: 9,39e-47 Length: 951
Score: 717.00 Matches: 149
Percent Similarity: 69.78% Conservative: 38
Best Local Similarity: 55.60% Mismatches: 45
Query Match: 41.42% Indels: 36
DB: 23 Gaps: 7

US-09-709-103-3fl (1-332) x AAS90571 (1-951)

oy 77 ArgMetValIleLeuGlySerIleValGlySerThrAlaIleValSerArgPheLeu 96

Db 142 CGCATGATGATGCTGGGCTCTCGGAGGAGAGCTTCATCTGCTCGCTTC 201
Qy 97 ThrGlyArgPheGluAspAlaTyrThrProThrIleGluAspPheHisArgGlySerPheTyr 116
Db 202 AATGGCCGCTTGAAGACACATACACACCCATCGAGAGACTTCCACCGTAAGATATAC 261
Qy 117 SerIleArgGlyGluValTyrGluLeuAspIleLeuAspThrSerGlyAsnHisProPhe 136
Db 262 AACATCCGCGGACATGATACAGCTGCACATCTCGATACCTTGGCAACACCCCTTC 321
Qy 137 ProAlaMetArgArgPheSerIleLeuThrGlyAspVal---PheIleLeuValPheSer 155
Db 322 CCCGCAATGGCGAGGCTGTGTCATCTCAGAGTAGGCGCCACTGCTGCTGGCTGGGAGC 381
Qy 156 LeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGln----- 170
Db 382 -----GGCAGGCGCAGGAGCATGGGTCGAGAGTGTG 411
Qy 171 -----GlnIleLeuAspThrIleSerCySLeuValAsnIleThrIleGluAsn 186
Db 412 CTGGCACTTAGCAGATCTCGAGGTCAAGTCTGCTGAGAGACCAAGAGAGCG 471
Qy 187 ValAspValProLeuValIleCySgIyAsnIyGlyAspArg---AspPheTyrArgIle 205
Db 472 GCGGAGCTGCCATGATCATCTGTGGCAGACAGACACGAGCGAGCTGTGGCGCAG 531
Qy 206 ValAspGlnArgGluIleGluGlnLeuValGlyAspAspProGlnArgCySalatTyrPhe 225
Db 532 GTGCCACCCAGGAGCGGAGCTGTGTGGGCGAC---GAGAACTGCGCTCACTTC 588
Qy 226 GluIleSerAlaIyAsnSerSerLeuAspGlnMetPheArgAlaLeuPheAlaMet 245
Db 589 GAGCTGTGGCCAGAGACACCAACCTGAGCGAGAGTGTCTACGCTCTTCACAGATG 648
Qy 246 AlaIyLeuProSerGluMetSerProAspLeuHisArgIyValSerValGlnTyrCyS 265
Db 649 GCCAGCTGCCACAGATGAGCCGCCCTGATGCAATCTCGTCCATACGCT 708
Qy 266 AspValLeuHisIleIyS---AlaLeuArgAsnIySLeuLeuArgAlaGlySer 283
Db 709 GAGCCCTTCACCCAGCGCTTCGATGACCCGCGTCAAGAGATG----- 756
Qy 284 GlyGlyGlyGlyAspProGlyAspAlaPheGlyIleValAlaProPheAlaArgArg 303
Db 757 -----GACGCTATGAGATGTCTGCCCTTCGCCCGCGC 792
Qy 304 ProSerValHisSerAspLeuMetTyrIleArgGluValAlaSerAlaGlySerGlnAla 323
Db 793 CCCAGCTCAACATGATCACTCAAGTACATCAAGCCAGAGTCTCTTGGAGAGCCAGCC 852
Qy 324 LysAspIySgIuArgCySValIle 331
Db 853 CGTGAAGAGACAGTGCACCATC 876

RESULT 15
ABLO7789 standard; cDNA; 1305 BP.
ID ABLO7789;
AC ABLO7789;
XX 26-MAR-2002 (first entry)
DT
XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 17849.
DE
XX Drosophila; developmental biology; cell signalling; insecticide;
KW Drosophila; gene; ss.
KW pharmacological; gene; ss.
OS Drosophila melanogaster.
XX MO2001.71042-A2.
FN 27-SEP-2001.
XX



GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 30, 2002, 18:04:26 ; Search time 2185 Seconds
(without alignments)
2460.825 Million cell updates/sec

Title: US-09-709-103-3f1
Perfect score: 1731

Sequence: 1 GIPGAGAPSPRAQPEQSP.....IREKASAGSQAKKRCVTS 332

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q/cgn2_1/USPTO.spool/SULLIVAN3f1/runat 30122002 144154 22149/app query.fasta_1.519
-DB=EST -QWMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosu62 -TRANS=human40.cdi -LIST=45
-LOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODES=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=SULLIVAN3f1 @CGN_1_1716 -runat 30122002 144154 22149 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST.*
1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_hic.*
9: gb_est1.*
10: gb_est2.*
11: gb_hic.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: gb_gss.*
18: em_gss_hum.*
19: em_gss_inv.*
20: em_gss_pln.*
21: em_gssprt.*
22: em_gss_fun.*
23: em_gssmam.*
24: em_gssmus.*
25: em_gss_othr.*
26: em_gss_pro.*
27: em_gss_rtd.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1482.5	85.6	1035	14	BM919341	BM919341 AGENCOURT
2	1437	83.0	962	13	BM543472	BM543472 AGENCOURT
3	1416	81.8	1103	14	BM920514	BM920514 AGENCOURT
4	1372.5	79.3	1053	14	BM921737	BM921737 AGENCOURT
5	1368	79.0	965	9	AL533318	AL533318 AL533318
6	1310	75.7	1023	13	BM543630	BM543630 AGENCOURT
7	1305	75.4	1032	14	BQ067637	BQ067637 AGENCOURT
8	1190.5	68.8	1137	14	BM921656	BM921656 AGENCOURT
9	1189	68.7	814	13	B1596688	B1596688 AGENCOURT
10	1186	68.5	742	10	AW028127	AW028127 W226C07.X
11	1168.5	67.5	1160	14	BM805574	BM805574 AGENCOURT
12	1162.5	67.2	831	13	B1596637	B1596637 603243203
13	1114.5	64.4	958	14	BQ719566	BQ719566 AGENCOURT
14	1097	63.4	699	12	BG706012	BG706012 602669154
15	1071	61.9	699	12	BG709229	BG709229 602674615
16	1071	61.9	700	13	B1596509	B1596509 603243558
17	1064	61.5	649	12	BG085090	BG085090 H3108B05-
18	1017	58.8	666	13	B1549939	B1549939 603194786
19	994.5	57.5	726	12	BF613135	BF613135 de30e03.Y
20	994	57.4	655	13	B1545172	B1545172 603242791
21	987	57.0	674	13	B17601563	B17601563 pglin.pk0
22	976.5	56.4	674	13	BM807669	BM807669 AGENCOURT
23	946.5	54.7	1438	14	B1553776	B1553776 603190722
24	926	53.5	758	13	BG969048	BG969048 602835003
25	924.5	53.4	644	13	BB636889	BB636889 B863889
26	920.5	53.2	648	10	BB632699	BB632699 B863889
27	920.5	53.2	689	10	BQ073742	BQ073742 AGENCOURT
28	888	51.3	1006	14	BQ073742	BQ073742 AGENCOURT
29	887	51.2	710	13	B1526038	B1526038 B1526038
30	869	50.2	506	13	AK015898	AK015898 Mus muscu
31	867	50.1	1300	11	BQ954076	BQ954076 AGENCOURT
32	866	50.0	947	14	BQ954076	BQ954076 AGENCOURT
33	862	49.8	556	12	BG654362	BG654362 B139604.Y
34	861.5	49.8	617	13	B1393669	B1393669 B139604.Y
35	858	49.6	904	14	BQ947936	BQ947936 AGENCOURT
36	853.5	49.3	1404	13	BM548645	BM548645 AGENCOURT
37	852	49.2	531	13	B1823596	B1823596 603040960
38	852	49.2	539	13	B1760810	B1760810 603040960
39	852	49.2	541	13	B1759201	B1759201 603042628
40	852	49.2	556	13	B1913105	B1913105 603179727
41	849	49.0	520	13	B1823533	B1823533 603041088
42	841	48.6	1332	13	BM460899	BM460899 AGENCOURT
43	837.5	48.4	640	13	B1490945	B1490945 B1490945
44	827	47.8	589	13	BM426066	BM426066 pgf2n.pk0
45	825	47.7	540	13	BI824314	BI824314 603040731

ALIGNMENTS

RESULT 1
BM919341
LOCUS
DEFINITION AGENCOURT 6715681 NIH_MGC_120 Homo sapiens CDNA clone IMAGE:5748585
5', mRNA sequence.
ACCESSION BM919341
VERSION BM919341.1 GI:19369720
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1035)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL
COMMENT

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at:
http://image.llnl.gov
Plate: L1AM12776 row: j column: 10
High quality sequence stop: 658.
Location/Qualifiers

FEATURES
source

1. 1035
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5748585"
/clone_lib="NIH MGC_120"
/lab_host="DH10B"
/note="Organ: pooled pancreas and spleen; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of spleen and pancreas from 28 yo male. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.5 kb, insert size range 1-2.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 025. Note: this is a NIH MGC Library."
BASE COUNT 209 a 386 c 293 g 146 t 1 others
ORIGIN

Alignment Scores: 4,75e-120 Length: 1035
Pred. No.: 1482.50 Matches: 293
Score: 95.81% Conservative: 8
Percent Similarity: 94.52% Mismatches: 5
Best Local Similarity: 85.64% Indels: 2
Query Match: 14 Gaps: 2
DB:

US-09-709-103-3f1 (1-332) x BM919341 (1-1035)

QY 3 ProSerGIYAAGIYAAProSerProSerAGIAGInProGInGInSerProProA 22
DB 40 CCGAGCGGAGCCGAGCCGAGCCGAGCCGAGCCGAGCCGAGCCGAGCCGAGCC 99
QY 23 AAlaHisProLaCyHisProSerAspProGInProLeuSerAlaLeuSerAlaPro 42
DB 100 GCTCACCCCGCGTGCACCCGAGCAGCCTTCAGCCGCTTCGCTTCGCGCCG 159
QY 43 ArpProProSerArpProLeuCySPrometLysLeuAlaAlaMetLysLysMetCyS 62
DB 160 CCGCCGCGCTCGCGGCCCTCTGCCAATGAATGCGCGCATGTATCAAGAAATGTC 219
QY 63 ProSerAspSerGIuLeuSerLLeProAlaLysAspCyTArGmetValLLeuGly 82
DB 220 CCGAGCAGCTCGGAGTATTCGCGCCAGAACTGCTATGCACTGCTATCTCGGC 279
QY 83 SerSerLysValGIyLysThrAlaLysValSerArpPheLeuThrGIyArpPheGIuasp 102
DB 280 TCGTCCAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 339
QY 103 AlaTyrThrProThrLLeGIuAspPheHisArGIyPheTyrSerLLeArGIyGIuVal 122
DB 340 GCTTACACGCGCTTACATCGAGACTTCACCGCAAGTTCTACTCATTCGCGCGAGGTC 399
QY 123 TyrGIuLeuAspLLeuAspThrSerGIyAsnHisProPheProAlaMetArGIuLeu 142
DB 400 TACAGCGTGAATCTCTGACAGCTTCGCGCAACACCGTTCCCGCGCATGCGCGCTTC 459
QY 143 SerLLeuThrGIyAspValPheLLeuValPheSerLLeuAspAsnArGIuAspSerPhe 162
DB 460 TCCATCTCCAGGAGAGCGTTTCTATCTGCTGATCGATCGCAACCGGAGCTCTTC 519

QY 163 GIuGIuValGIuArGIuLeuArGIuGInLLeuLLeuAspThrLysSerCySLeuLysAsnLys 182
DB 520 GAGAGAGTCAGAGGCTCAGGAGCAGATCTTCAGACACAAAGTCTTCAGAAACAA 579
QY 183 ThrLysGIuAsnValAspValProLeuValLLeCySGIyAsnLysGIyAspArGIuAspPhe 202
DB 580 ACCAAGAGAGAGCTGAGCTGCTCCCTGTCATCTCGGCAACAGGGTGAACCGGAGCTTC 639
QY 203 TyrArGIuValAspGIuArGIuLLeGIuLLeuValGIyAspAspProGIuArGIyCys 222
DB 640 TACCGAGAGTGAACAGCGGAGATCGAGCTGTGTGGGAGAGACCCAGCGCTGC 699
QY 223 AlaTyrPheGInLLeSerLAlaLysLysAsnSerSerLLeuAspGIuMetPheArGIuLLeu 242
DB 700 GCTTACTTCAGATCTTCGCGCAAGAAACAGAGCTGAGACCAATGTTCCGCGCTTC 759
QY 243 PheAlaMetAlaLysLeuProSerGIuMetSerProAspLLeuHisArGIyValSerVal 262
DB 760 TTCGCCATGGCCAGAGCTGCCAGAGATGAGCCAGACTCGACCGCAAGGTTCTCGGTG 819
QY 263 GIuTyrCySAspValLeuHisLysLysValAlaLeuArGIuAsnLysLysLeuArGIuLLeu 282
DB 820 CAGTACTGCGACCTGCTGCAACAGAAAGCGCTGCGGAAACAAAGAACTGCTGCGGCGCGC 879
QY 283 SerGIy-GIyGIy-GIyGIyAspProGIyAspAlaPheGIyLLeuAlaLLeu-----Prop 300
DB 880 AACG 939
QY 300 heaLArGIy--ArpProSerValHis 307
DB 940 TCG 965

RESULT 2
BM543472 962 bp mRNA linear EST 20-FEB-2002
LOCUS
DEFINITION
AGENCOURT 6492614 NIH_MGC_124 Homo sapiens CDNA clone IMAGE:5726554
5', mRNA sequence.
BM543472
BM543472.1 GI:18773895
VERSION
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 962)
NIH-MGC http://mgc.nci.nih.gov/.
TITLE
JOURNAL
COMMENT
Email: cgaabs-remail.nih.gov
Tissue Procurement: Invitrogen
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at:
http://image.llnl.gov
Plate: L1AM12719 row: d column: 11
High quality sequence start: 18
High quality sequence stop: 692.
Location/Qualifiers

FEATURES
source

1. 962
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5726554"
/clone_lib="NIH MGC_124"
/tissue_type="hippocampus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: EcoRV (destroyed); Site_2: NotI; RNA source male hippocampus, age 27. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 0.9-4 kb. Library is

3	ProSerGlyAlaGlyAlaProSerProSerArgAlaGlnProGluGlnSerProProAla	22
b	54	CCGAGCGGAGCGGAGGCCCAAGCCGAGCGCGCCGACGCCGAGCAGCCCTTCAGCC
113	23	AlaHisProAlaCysHisProSerAspProGlnProLeuSerAlaLeuLeuSerAlaPro
42	114	GCTCACCGCGGTGCACGCCAGGACCTCAGCGGCTCTTCGCCCTTCTCGGCCCGC
173	43	ArgProProSerArgProLeuCysProMetLysLeuAlaAlaMetIleLysLysMetCys
62		

```

Db 174 CCGCCGCCCTCCGGGCCCCCTGCCCCAATGAACCTGGCGGAGATCAACAAAGATGTC 233
Qy 63 ProSerAspSerGluLeuSerIleProAlaLysAsnCyTYrArgMetValIleLeuGly 82
Db 234 CCGAGGAGCTCGAGAGCTGATCCGGGCCAAGAACTGTAATCGCATGTCATCTCGGC 293
Qy 83 SerSerIleValGlyIleThrAlaIleValSerArgPheLeuThrGlyArgPheGluAsp 102
Db 294 TCGTCAAGGTGGCAAGACCGGCATCGTGGCGCTTCTCCACCGCGCGCTTCAGAGAC 353
Qy 103 AlaIleThrProThrIleGluAspPheHISArgGlySerIleArgGlyGluVal 122
Db 354 GCTACACGCGCTACATCCAGAGCTTCACCGCAAGCTTCTTACTCATCCGCGGAGGTC 413
Qy 123 TyrGluLeuAspIleLeuAspThrSerGlyAsnHISProPheProAlaMetArgArgLeu 142
Db 414 TACCAAGCTCGACATCTCCGACACGTCGGCAACACCGCTCCCGCCATCGCGGCGCTC 473
Qy 143 SerIleLeuThrGlyAspValPheIleLeuValPheSerIleAspAsnArgAspSerPhe 162
Db 474 TCCATCTCCACAGAGAGAGCTTTCATCTGTCATCTGTCAGTGGACAACTGATCTCTTC 533
Qy 163 GluGluValGlnArgLeuArgGlnGlnIleLeuAspThrIleSerCysLeuLysAsnLys 182
Db 534 GAGAGAGTGACGCGCTCAGGACACATCTCGACACCAAGCTTGCTCAAGAACAA 593
Qy 183 ThrIleGluAsnValAspValProLeuValIleCysGlyAsnLysGlyAspArgAspPhe 202
Db 594 ACCAAGAGAAAGCTGACGTCGCTCCGTCATCTGCGGCAACAAAGGTGACCGCGACTTC 653
Qy 203 TyrArgGluValAspGlnArgGlnIleGluGlnLeuValGlyAspAspProGlnArgCys 222
Db 654 TACCCGCGAGTGACACCGGAGATCGACAGCTGTCGTCGCGACGACCCCGACGCTGC 713
Qy 223 AlaIlePheGluIleSerAlaLysLysAsnSerSerLeuAspGlnMetPheArgAlaLeu 242
Db 714 GCTTACTTGGAGATCTCGGCGCAAGAAAG-AGCAGCTCGGACCAAGATGTTCCGCGCTC 772
Qy 243 PheAlaMetAlaLysLeu-ProSerGluMetSerProAspLeuHISArgLysValSerVa 262
Db 773 TTCCGATCGGCGAGCTGCGCGGAGATGAGCCCAAGCTCGACCGCAAGGTCTGGGT 832
Qy 262 IGHTrCyAspValIleHIS-LysLysAlaLeuArgAsnLys-LysLeuLeuArgAla 281
Db 833 GCGATCTCGAGAGCTCTCCACAAAGAGGCTTCGGAACAAAGAACTCTTCGCGGC 892
Qy 282 GlySerGlyGlyGlyGlyAspProGlyAspAla-----PheGlyIleValAlaPro 299
Db 893 CCGGCAACCGCGCGCGCGCGCGCGCAACCGGCGCAACCTTTGGCATCTCGGAGCC 952
Qy 300 Phe---AlaArgArgProSerValHIS-SerAspLeuMetTyrlleArgGlyLysAla 318
Db 953 TTTCGCGCGCGCGCGCGCGCGCGCGCAAGGTACAGGAACCTTAAGAAATCCCGCAAAAGCCA 1012
Qy 318 r---AlaGlySerGlnAlaLysAspLys 326
Db 1013 ACCGCCGCGNACGCGCAAGCAAGAAAA 1040

RESULT 4
BM921737 1053 bp mRNA linear EST 12-MAR-2002
LOCUS BM921737
AGENCY NCUR 6708101 NIH_MGC_115 Homo sapiens CDNA IMAGE:5753301
DEFINITION 5', mRNA sequence.
ACCESSION BM921737
VERSION BM921737
KEYWORDS EST.
SOURCE Homo sapiens
ORGANISM human.
REFERENCE 1 (bases 1 to 1053)
AUTHORS Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgi.nci.nih.gov/.

```

```

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
CONTACT Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgabs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L14M12788 row: n column: 22
High quality sequence stop: 600.
FEATURES
source
location/Qualifiers
1..1053
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5753301"
/clone_id="NIH_MGC_115"
/lab_host="MDH10B"
/note="Organ: pooled brain, lung, testis; Vector:
PCMV-SpOrie; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb.
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH-MGC Library."
BASE COUNT 215 a 390 c 297 g 150 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 2,176-110 Length: 1053
Score: 1372.50 Matches: 277
Percent Similarity: 90.10% Conservative: 5
Best Local Similarity: 88.50% Mismatches: 22
Query Match: 79.29% Indels: 9
DB: 14 Gaps: 3
US-09-709-103-3fl (1-332) x BM921737 (1-1053)
Qy 3 ProSerGlyAlaGlyAlaProSerProSerArgAlaGlnProGluGlnSerProProAla 22
Db 53 CCGAGCGGAGCGGAGCGGCCCAAGCCCGAGCGCGGCCCGCAAGAGAGAGATGTC 112
Qy 23 AlaHISProAlaCySHISProSerAspProGlnProLeuSerAlaLeuSerAlaPro 42
Db 113 GCTCAACCGCGGTGCACCCAGAGACCTTACGCGCTCTCTGCGCTTCTCGCGCCG 172
Qy 43 ArgProProSerArgProLeuCySPrometLysLeuAlaMetIleLysLysMetCys 62
Db 173 CCGCCGCCCTCCGGGCCCCCTGCCCCAATGAACCTGGCGGAGATCAACAAAGATGTC 232
Qy 63 ProSerAspSerGluLeuSerIleProAlaLysAsnCyTYrArgMetValIleLeuGly 82
Db 233 CCGAGGAGCTCGAGAGCTGATCCGGGCCAAGAACTGTAATCGCATGTCATCTCGGC 292
Qy 83 SerSerIleValGlyIleThrAlaIleValSerArgPheLeuThrGlyArgPheGluAsp 102
Db 294 TCGTCAAGGTGGCAAGACCGGCATCGTGGCGCTTCTCCACCGCGCGCTTCAGAGAC 352
Qy 103 AlaIleThrProThrIleGluAspPheHISArgGlySerIleArgGlyGluVal 122
Db 354 GCTACACGCGCTACATCCAGAGCTTCACCGCAAGCTTCTTACTCATCCGCGGAGGTC 412
Qy 123 TyrGluLeuAspIleLeuAspThrSerGlyAsnHISProPheProAlaMetArgArgLeu 142
Db 414 TACCAAGCTCGACATCTCCGACACGTCGGCAACACCGCTCCCGCCATCGCGGCGCTC 472
Qy 143 SerIleLeuThrGlyAspValPheIleLeuValPheSerIleAspAsnArgAspSerPhe 162

```

```

Db 473 TCCATCTCTCAGGAGACGCTTTTCATCTCTGGTGTTCAGTCTGGACACCGGACTCCTTC 532
Qy 163 GluGluValGlnArgLeuArgGlnGlnLeuAspThrLysSerCysLeuLysAsnLys 182
Db 533 GAGGAGTGCAGGCTCAGGAGCAGATCTCTGACACCAAGTCTTGCCTCAAGAACAAA 592
Qy 183 ThrLysGluAsnValAspValProLeuValLysCysGlyAsnLysGlyAspArgAspPhe 202
Db 593 ACCAAGAGAACGTCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 652
Qy 203 TyrArgGluValAspGlnArgGlnLeuValLysCysGlyAsnLysGlyAspArgAspPhe 222
Db 653 TACCGGAGGTGACACCGGAGATCGAGCAGTGTGGCGACGACCCAGCGCTGC 712
Qy 223 AlaTyrPheGluLysSerAlaLysAsnSerSerLeuAspGlnMetPheArgAlaLeu 242
Db 713 GCTTACTTTCGAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 772
Qy 242 uPheAlaMetAlaLysLeuProSerGluMetSerProAspLeuHisArgLys--Valse 261
Db 773 CTGCGCATGCTCCAGTGTCCAGCGAGATGAGCCGAGACCTGCACCGCAAGGTCCTCG 832
Qy 261 rValGlnTyrCysAspValLeuHisLysLys---AlaLeuArgAsnLysLysLeuLeuAr 280
Db 833 GCGAAGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 892
Qy 280 g-AlaGlySer-----GlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 295
Db 893 GGGCGGTATACCGTGTGGCGGCGCGGAGGAGACCGGCGGAGACCCCTTTGGGGAATCG 952
Qy 295 lylleValAlaPropheAlaAlaArgArgProSerVal 306
Db 953 GGGGCAACCTTCTGTGGGCGCGGAGACCAAGGTA 987

RESULT 5
AL533318 965 bp mRNA linear EST 13-FEB-2001
LOCUS AL533318 LTI_FLO15_Brn1 Homo sapiens cDNA clone CS0DN003YJ19 5
DEFINITION AL533318 LTI_FLO15_Brn1 Homo sapiens cDNA clone CS0DN003YJ19 5
ACCESSION AL533318
VERSION AL533318.1 GI:12796811
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Li.W.B., Gruber.C., Jesse,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers
1. 965
/organism="Homo sapiens"
/db-xref="taxon:9606"
/clone="CS0DN003YJ19"
/clone_lib="LTI_FLO15_Brn1"
/sex="male"
/tissue_type="Adult brain"
/notes="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was constructed by life Technologies.
Contact : Feng Liang Life Technologies, a division of
Invitrogen 9800 Medical Center Drive Rockville, Maryland
20850, USA Fax : (1) 301 610 8371 Email :
flying@lifetech.com URL :
http://fulllength.invitrogen.com"
189 a 352 c 273 g 144 t 7 others

BASE COUNT

```

ORIGIN

```

Alignment Scores:
Score: 4.83e-110 Length: 965
Pred. No.: 1368.00 Matches: 282
Percent Similarity: 95.00% Conservative: 3
Best Local Similarity: 94.00% Mismatches: 14
Query Match: 79.03% Indels: 5
DB: 9 Gaps: 1

US-09-709-103-3P1 (1-332) x AL533318 (1-965)

Qy 3 ProSerGlyAlaGlyAlaProSerProSerArgAlaGlnProGlnSerProAla 22
Db 44 CCGGAGCGGAGGAGCCCAAGCCCGGAGCCCGGAGCAGAGCCCTCCAGCC 103
Qy 23 AlaHisProAlaCysHisProSerAspProGlnProLeuSerAlaLeuLeuSerAlaPro 42
Db 104 GCTCACCCCGGTGCCACCCCGGAGCCCTCAGCCGCTCTCTGCCCCCTCTCTCGGCCCTG 163
Qy 43 ArgProProSerArgProLeuCysProMetLysLeuAlaAlaMetLysLysMetCys 62
Db 164 CGCCCGCCCTCGCGGCCCTCTGCCCAATGAACTGGCCGCGATGATCAAGAAGATGTC 223
Qy 63 ProSerAspSerGluLeuSerIleProAlaLysAsnCysTyrArgMetValIleLeuGly 82
Db 224 CCGAGCGACTCGGAGCTGAGTATCCCGGCCCAAGACTGCTATCGCATGGTCACTCTCGGC 283
Qy 83 SerSerLysValGlyLysThrAlaIleValSerArgPheLeuThrGlyArgPheGluAsp 102
Db 284 TCGTCCAGGTGGGCAAGACGGCCATCGTGTGCGCTTCTCACCGGCGCTTCGAGGAC 343
Qy 103 AlaTyrThrProThrIleGluAspPheHisArgLysPheTyrSerIleArgGlyGluVal 122
Db 344 GCCTACAGCTACCATCGAGACTTCCACCGCAAGTCTACTCATCCGCGGCGAGGTC 403
Qy 123 TyrGlnLeuAspIleLeuAspThrSerGlyAsnHisProPheProAlaMetArgLeu 142
Db 404 TACCAGCTCGACATCTCGACAGCTCCCGCAACCAACCCGCTTCCCGCGCATCGC-CGCTC 462
Qy 143 SerIleLeuThrGlyAspValPheIleLeuValPheSerLeuAspAsnArgAspSerPhe 162
Db 463 TCCATCTCACAGGAGAGCTTTTCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTTTC 522
Qy 163 GluGluValGlnArgLeuArgGlnIleLeuAspThrLysSerCysLeuLysAsnLys 182
Db 523 GAGGAGTGCAGCGGCTCAGGAGCAGATCTCGACACCAAGTCTTGGACACCGGACTCTTC 582
Qy 183 ThrLysGluAsnValAspValProLeuValLysCysGlyAsnLysGlyAspArgAspPhe 202
Db 583 ACCAAGAGAGACGTGGAGCTGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 642
Qy 203 TyrArgGluValAspGlnArgGlnLeuValLysCysGlyAsnLysGlyAspProGlnArgCys 222
Db 643 TACCGGAGGTGGACCGCGGAGATCGAGCAGTGTGGCGGAGCAGCCCGGAGCGCTGC 702
Qy 223 AlaTyrPheGluLysSerAlaLysLysAsnSerSerLeuAspGlnMetPheArgAlaLeu 242
Db 703 GCTTACTTTCGAGATCTCGGCCAAGAGAGAGAGCGCTGGACCATGTTCTCGCGCGCTC 762
Qy 243 PheAlaMetAlaLysLeuProSerGluMetSerProAspLeuHisArgLysValSerVal 262
Db 763 TTCGCCATGCGCAAGTGTGCCCGAGATGAGCCGAGACTCTGACCGCAAGTCT-TCGGTG 821
Qy 263 GlnTyrCysAspValLeuHisLysLysAlaLeuArgAsnLysLysLeuLeuArgAlaGly 282
Db 822 CAGTACTGCGAGCTGTGCAAAAGAGCGCTGCGGAGCAA-AARCTGCTGCGGCGGCA 880
Qy 283 SerGlyGlyGlyGlyAspProGlyAspAlaPheGlyIleValAlaProPheAlaArg 302
Db 881 GCGS--GGCGGCGCGGAGCGGCG--AGCTTTTGCATCGTGGCACCTTGGCGCGCG 935

RESULT 6

```

BM543630 1023 bp mRNA linear EST 20-PEB-2002
 LOCUS BMS43630 6492527 NIH_MGC_124 Homo sapiens cDNA clone IMAGE:5726764
 DEFINITION 5', mRNA sequence.
 ACCESSION BMS43630
 VERSION BMS43630.1 GI:18774186
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1023)
 NIH-MGC http://mgs.nci.nih.gov/.
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished (1999)
 JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT Email: cgabbs-remail.nih.gov
 Tissue Procurement: Invitrogen
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LLM12719 row: m column: 05
 high quality sequence stop: 637.
 Location/Qualifiers
 1..1023
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5726764"
 /clone_1ib="NIH_MGC_124"
 /tissue_type="hippocampus"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pCMV-Sport6; Site: 1: EcorV
 (destroyed); Site 2: NotI; RNA source male hippocampus,
 age 27. Library is oligo-dT primed and directionally
 cloned (EcorV site is destroyed upon cloning). Average
 insert size 1.4 kb, insert size range 0.9-4 kb. Library is
 normalized and enriched for full-length clones and was
 constructed by C. Gruber (Invitrogen). Research Genetics
 tracking code 012."
 BASE COUNT 196 a 378 c 281 g 168 t
 ORIGIN
 Alignment Scores:
 pred. No.: 6.36e-105 Length: 1023
 Score: 1310.00 Matches: 271
 Percent Similarity: 85.98% Conserved: 5
 Best Local Similarity: 84.42% Mismatches: 30
 Query Match: 75.68% Indels: 15
 DB: 13 Gaps: 3
 US-09-709-103-3f1 (1-332) x BMS43630 (1-1023)
 QY 3 ProSeGjYlAglYAlAProSeRProSeRArgAlAGlProGluGlnSerProProAla 22
 Db 2 CCGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 61
 QY 23 AlAhIsProAlaCyShIsProSeRArgProGlnProLeuSerAlaLeuSerAlaPro 42
 Db 62 GCTACACCGCGTCCAGCCAGCGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 121
 QY 43 ArgProProSeRArgProLeuCyProMeLysLeuAlaLeuSerAlaLeuSerAlaPro 62
 Db 122 GCGCGCGCTCGCGGCGGCTTCCATGAACTGGCGCGGAGTGTATCAAGAGATGTC 181
 QY 63 ProSeRArgSeRgLeuSerAlaProAlaLysAsnYerYrArgMetValIleLeuGly 82
 Db 182 CCGAGCGAGCTCGAGCTAGTATCCCGGCAAGAACTGCTATGCGATGTCATCCCGG 241
 QY 83 SerSerLygValGlyLysThrAlaIleValSerArgPheLeuThrgLyArgPheGluAap 102

Db 242 TCGTCAAGGTGGCAAGACGCGCATCGTTCGCGCTTCTCACCAGCGCGCTTGGAGAC 301
 QY 103 AlAtYrThProThrlLeglUaSPheHIsArgLysPheYrSerIleArgGlyGluVal 122
 Db 302 GCCTACACGCTTCCATCGAGAGACTTCCACCGCAAGTTCTTATCTCAGCGGAGGTC 361
 QY 123 TyrGlnLeuAspIleLeuAspThrSerGlyAsnHIsProPheProAlaMetArgArgLeu 142
 Db 362 TACCAAGCTCGACATCTTCGACACGTCGCGCAACACCGCTTCCCGGCAATGGCGGCTTC 421
 QY 143 SerIleLeuThrgLyAspValPheIleLeuValPheSerLeuAspAsnArgAspSerPhe 162
 Db 422 TCCATCTCCACAGAGACGTTTTCATCTCTGTTCTGTTGACACACCGCACTTCTTC 481
 QY 163 GluGluValGlnArgLeuArgGlnIleLeuAspThrLysSerCyLeuLysAsnLys 182
 Db 482 GAGAGGTGACGCGGCTCAGGACAGCATCTCGACACCAAGCTTTCCTCAAGAACAA 541
 QY 183 ThrLygLeuAsnValAspValProLeuValIleCySgLyAsnLySgLyAspArgAspPhe 202
 Db 542 ACCAAGAGAGAACGTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 601
 QY 203 TyrArgGluValAspGlnArgGlnIleGluGlnLeuValGlyAspAspProGlnArgCyS 222
 Db 602 TACCGGAGGTGACGAGCGGAGATCGAGCGGTGCGGAGCGGAGCGGAGCGGAGCTTC 661
 QY 223 AlAtYrPheGluIleSerAlaLysLysAsnSerSerLeu-AspGlnMetPheArgAlaLe 242
 Db 662 GCTTACTTCGAGATCTCGGCGCAAGAAACACACCTTGGGACAGATGTTCCGCGGCT 721
 QY 242 UpheAlaMetAla-LysLeu-ProSeRgLeuMetSerProAspLeuHIsArgLys---Val 260
 Db 722 CTTTCCATGCGGCGGAGCTGCGGCGGAGATGAGACCGGAGCTGCGGAGGTC 781
 QY 261 SerValGlnTyrCyS---AspValLeuHIsLysLysAlaLeuArgAsnLysLeuLeu 279
 Db 782 GGTGAGGATGATCGGCGAGCGGCTGCGCAAGAAAGCGCTGCGGAGAACACAGAACTG 841
 QY 280 ArgAlaGlySerGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 289
 Db 842 GCTGCGGAGCGGCGGAGATCGGCGGCGGCTTCTTATGATTAACCTTGACCGAGAGC 901
 QY 290 ProGlyAspAlaPheGlyIleValAlaProPheAlaArgArgProSerValHisSer 308
 Db 902 CTTGGGCGAGCTGAGCGGCGGAGCTTTCGCGACAGTCCGCGCTTCTTGTGATCC 958
 RESULT 7
 B0067637 1032 bp mRNA linear EST 02-APR-2002
 LOCUS B0067637 6759053 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5755214
 DEFINITION 5', mRNA sequence.
 ACCESSION B0067637
 VERSION B0067637.1 GI:19896683
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1032)
 NIH-MGC http://mgs.nci.nih.gov/.
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished (1999)
 JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT Email: cgabbs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LLM12793 row: n column: 15
 high quality sequence stop: 602.

QY 63 ProSerAspSerGluLeuSerIleProAlaLysAsnCysTyrArgMetValIleGly 82
 Db 253 CCGAGCGAGTCCGAGCTAGTATCCGCGCAAGAACTCTATCGCATGGTCTCTCGGC 312
 QY 83 SerSerLysValGlyLysThrAlaIleValSerArgPheLeuThrGlyArgPheGluAsp 102
 Db 313 TCGTCCAAAGTGGGCAAGACGCCATCGTGTGCGCTTCTCACCAGCGGCTTCGAGGAC 372
 QY 103 AlaTyrThrProThrIleGluAspPheHisArgLysPheTyrSerIleArgGlyVal 122
 Db 373 GCTTACACGCTTACCATCGAGGACTTCCACCGCAAGTTCTACTCCATCCGCGGCGAGTC 432
 QY 123 TyrGlnLeuAspIleLeuAspThrSerGlyAsnHisProPheProAlaMetArgArgLeu 142
 Db 433 TACCAGCTGCATCTCGACAGTCCGCGCAACACCGCTTCCCGCGCATCGGCGCTC 492
 QY 143 SerIleLeuThrGlyAspValPheIleLeuValPheSerLeuAspAsnArgSerPhe 162
 Db 493 TCAATCTCTACAGGAGAGCTTTTCACTCTGTTGTTCACTGAGTCTGGACAAACCGGACTCTTC 552
 QY 163 GluGluValGlnArgLeuArgGlnIleLeuAspThrLysSerCysLeuLysAsnLys 182
 Db 553 GAGGAGTGCAGCGCTCAGGAGCAGATCTCTGACACCAAGTCTTCCCTCAAGAACAA 612
 QY 183 ThrLysGluAsnValAspValProLeuValIleCysGlyAsnLysGlyAspArgAspPhe 202
 Db 613 ACCAAGGAGAACGTGGAGCTGCCCTGTGTATCTGCGGCAACAAAGGTGACCGGACTTC 672
 QY 203 TyrArgGluValAspGlnArgGluIleGlnLeuValGlyAspAsp-ProGlnArgCy 222
 Db 673 TACCGGAGGTGGACCGCGGAGATCGAGCAGCTGTGTGGGGCGAGACCCCGCGGCTG 732
 QY 222 sAlaTyrPheGlu-IleSerAlaLysLysAsnSerSerLeuAspGlnMetPheArgAla 242
 Db 733 CGCTTACTTCAGAAATCTCGGCAAGAGAACACGACCGG-GACCAGATGTC-CGCGCGC 790
 QY 242 euPheAlaMetAlaLysLeu 248
 Db 791 TCTTCGCCCCATGGCAAGCTG 810

RESULT 10
 AW028127
 LOCUS
 DEFINITION w26c07.x1 NCI_CGAP Kid11 Homo sapiens cDNA clone IMAGE:2530668 3', similar to TR:035626 035626 RAS, DEXAMETHASONE-INDUCED 1 ;, mRNA sequence.
 ACCESSION AW028127
 VERSION EST.
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 742)
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D.
 DNA Sequencing by: Greg Lennon, Ph.D.
 Clone Distribution: NCI-CGAP University Genome Sequencing Center
 found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html
 Seq primer: -40UP from Gibco
 High quality sequence stop: 445.
 Location/Qualifiers
 1. .742

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2530668"
 /clone_lib="NCI_CGAP_Kid11"
 /lab_host="DH10B"
 /note="Organ: kidney; Vector: p773D-Pac (Pharmacia) with a modified polylinker; Site_1: Not 1; Site 2: Eco RI; Plasmid DNA from the normalized library NCI_CGAP Kid3 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (clones IDs 1322376-1323911, 1456007-1456775, and 1500552-1502855). Subtraction by Bento Soares and M. Fatima Bonaldo. "

BASE COUNT 149 a 260 c 208 g 122 t 3 others
 ORIGIN

Alignment Scores:
 Pred. No.: 3,25e-94 Length: 742
 Score: 1186.00 Matches: 231
 Percent Similarity: 95.93% Conservative: 5
 Best Local Similarity: 93.90% Mismatches: 10
 Query Match: 68.52% Indels: 0
 DB: 10 Gaps: 0

US-09-709-103-3F1 (1-332) x AW028127 (1-742)

QY 41 AlaProArgProSerArgProLeuCysProMetLysLeuAlaMetIleLysLys 60
 Db 2 GCCCGCGCGCGCCCTCGCGGCCCTCTGCCNATGAACTGGCCGCGATGATCAAGAG 61
 QY 61 MetCysProSerAspSerGluLeuSerIleProAlaLysAsnCysTyrArgMetValIle 80
 Db 62 ATGTCGCCGAGCGACATCGGAGCTGAGTATCCGCGGCAAGAACTGCTATCGATGTCATC 121
 QY 81 LeuGlySerSerLysValGlyLysThrAlaIleValSerArgPheLeuThrGlyArgPhe 100
 Db 122 CTGGCTCTGTCGAAGTGGGCAAGACGCGCATCGTGTGCGCTTCTCCTGCGCGCTTC 181
 QY 101 GluAspAlaTyrThrProThrIleGluAspPheHisArgLysPheTyrSerIleArgGly 120
 Db 182 GAGGACGCTTACACGCTACCATCGAGGACTTCCACCGCAAGTCTTACTCCATCCGCGGC 241
 QY 121 GluValTyrGlnLeuAspIleLeuAspThrSerGlyAsnHisProPheProAlaMetArg 140
 Db 242 GAGTCTTACCAGCTGCATCTCTGACACGCTCCGCAACACCCGCTTCCCGCCATGCGG 301
 QY 141 ArgLeuSerIleLeuThrGlyAspValPheIleLeuValPheSerLeuAspAsnArgAsp 160
 Db 302 CGCTCTCCATCTCACAGGAGACGTTTTTCATCTGTTGTTTCAGTCTGGCAACCGCGAC 361
 QY 161 SerPheGluGluValGlnArgLeuArgGlnIleLeuAspThrLysSerCysLeuLys 180
 Db 362 TCTTTCGAGGAGTGCAGCGCTCAGGCGAGAGATCTCCGACCAAGTCTTGTCTCAAG 421
 QY 181 AsnLysThrLysGluAsnValAspValProLeuValIleCysGlyAsnLysGlyAspArg 200
 Db 422 AACAAACCAAGAGAACGTGACGTGCCCTGCTGTCATCTCGGCAACCAAGGGTGACGC 481
 QY 201 AspPheTyrArgGluValAspGlnArgGluIleGluGlnLeuValGlyAspAspProGln 220
 Db 482 GACTTCTACCGGAGTGTGACCGCGCGAGATCGAGCAGCTGTTGGGCGACACCCCGCAG 541
 QY 221 ArgCysAlaTyrPheGluIleSerAlaLysLysAsnSerSerLeuAspGlnMetPheArg 240
 Db 542 CGTGTGCGCTTCTTCCAGATCTTCGCAAGAGAACACGACGCTGAGACCATGTTCCGC 601
 QY 241 AlaLeuPheAlaMetAlaLysLeuProSerGluMetSerProAspLeuHisArgLysVal 260
 Db 602 GGCCTCTTCGCGTCCGAGCTGCCAGCGAGATGAGCCAGACCTGCACCCGCAAGGTC 661
 QY 261 SerValGlnTyrCysAspValLeuHisLysLysAlaLeuArgAsnLysLysLeuLeuArg 280

```

Db 662 TCNGTGCACTGCTGACGCTGCTCCACAGAGCGCTGCCGAACAAGAGCTGTGCGG 721
Qy 281 AAGlySerGlyGlyGly 286
Db 722 TCCGACAGCGGCGCTGGT 739

RESULT 11
BM805574 1160 bp mRNA linear EST 05-MAR-2002
LOCUS BM805574
DEFINITION AGENCOURT 6498388 NIH_MGC_124 Homo sapiens cDNA clone IMAGE:5729129
5', mRNA sequence.
ACCESSION BM805574
VERSION BM805574.1 GI:19122397
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1160)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: Invitrogen
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM12725 row: o column: 18
High quality sequence start: 23
High quality sequence stop: 501.
Location/Qualifiers
1. 1160
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5729129"
/clone_11b="NIH_MGC_124"
/tissue_type="hippocampus"
/lab_host="DH10B"
/notes="Organ: brain; Vector: pCMV-SPORT6; Site: 1: EcoRV
(destroyed); Site 2: NotI; RNA source male hippocampus,
age 27. Library is oligo-dT primed and directionally
cloned (EcoRV site is destroyed upon cloning). Average
insert size 1.4 kb. Insert size range 0.9-4 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (Invitrogen). Research Genetics
tracking code 012."
BASE COUNT 207 a 473 c 279 g 201 t
ORIGIN

Alignment Scores:
Pred. No.: 1.92e-92 Length: 1160
Score: 1168.50 Matches: 229
Percent Similarity: 94.55% Conservative: 14
Best Local Similarity: 89.11% Mismatches: 10
Query Match: 67.50% Indels: 4
DB: 14 Gaps: 1

US-09-709-103-3f1 (1-332) x BM805574 (1-1160)
Qy 3 ProSerGlyAlaGlyAlaProSerProSerArgAlaGlnProGlnGlnSerProProAla 22
Db 50 CCGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 109
Qy 23 AlaHisProAlaGlySerHisProSerArgProGlnProLeuSerAlaLeuSerAlaPro 42
Db 110 GCTACCGCGGCTGCGACCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 169
Qy 43 ArgProSerArgProLeuCyseProMetLeuAlaMetIleGlyLeuMetCys 62

FEATURES
source
1. 831

```



```

Db      61 GTGGGCAAGACGGCCATCGTGTGCGCGCTTCTGACCGCGCGCTTCGAGAGACGCTACAGC 120
Qy      106 ProthrilleguaspPheH1sArgLysPheTySer11leargLylgluVal1TyrgInleu 125
Db      121 CTRACATTCAGAGACTTCCACCGCAAGTCTACCTCCATCCGCGCGAGGTCTACAGCTC 180
Qy      126 Asp11leuaspThrserGlyasnH1sProPheProAlaMetArgArgLeuSer11leu 145
Db      181 GACATCTCGACAGCTCCGGACACCCCGTCCCGCCATCGCGCCCTCTCCATCTC 240
Qy      146 ThrGlyaspValPhe1leuValPheSerLeuaspAsnArgAspSerPhegluVal 165
Db      241 ACAGAGACAGCTTTCATCTGTTGTTGATCTGACACACCGGACTCTTCGAGAGGTG 300
Qy      166 GlnArgLeuArgGlnGln11leuaspThrLysSerCysleuLysasnLysThrLysGlu 185
Db      301 CAGGGGCTCAGGACAGATCTCGACACCAAGTCTTGCCCAAGAACCAAGAGAG 360
Qy      186 AsnValaspValProleuVal11leCysglYasnLysglYaspArgAspPheTyArgGlu 205
Db      361 AACGTGACGTGCCCTGTGATCTTCCGCAACAAAGGTGACCCGACTTCTACCGGAG 420
Qy      206 ValaspGlnArgGln11leGluGlnleuVal1glYaspAspProGlnArgCysAlaTyPhe 225
Db      421 GTGACACAGCGGAGATCGAGAGCTGGTGGGCGACGACCCCGAGCGCTGCGCTACTTC 480
Qy      226 Gln11leSerAlaLysLysasnSerSerLeuaspGlnMetPheArgAlaLeuPheAlaMet 245
Db      481 GAGATCTCGCGCCAGAACAGACGCTGACCGAGATGTTCCGGCGCTCTTCGCCATG 540
Qy      246 AlaLysleuProSerGlnMetSerProaspLeuH1sArgLysValSerValGlnTyCys 265
Db      541 GCGAAGCTGCCAGAGATGAGCCAGACCTGCCCGGAGAGTCTCGATGAGAAATGC 600
Qy      266 AspVal1leuH1sLysLysAlaLeuArgasnLysLysleuLeuArg-AlaGlySerGly-G 285
Db      601 GACGTGCTCACAAGAGAGCGCTGCGGAAACAAAGCTGCTCGGCGCGGACGAGAG 660
Qy      285 lYglYglYglYaspPro---GlyaspAlaPheGly11leValAlaProPheAlaArg 303
Db      661 GCCGCGCGCTGACGCCCGGGGAGACCTTTTGAGATCGGGGGCCCCCTTTCGGGGCG 719

```

```

RESULT 14      699 bp      mRNA      linear      EST 07-MAY-2001
LOCUS      BG706012      602669154F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4792069 5',
DEFINITION      mRNA sequence.
ACCESSION      BG706012
VERSION      BG706012.1 GI:13980932
KEYWORDS
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

```

```

REFERENCE      1 (bases 1 to 699)
AUTHORS      NIH-MGC http://mgi.nci.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
Email: cgaabs-r@mail.nih.gov
Tissue Procurement: Miklos Palcovits, M.D., Ph.D. (NHGRI), Shitaki
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shitaki
Toshitaki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10669 row: K column: 14
High quality sequence stop: 696.
Location/Qualifiers
1..699
/organism="Homo sapiens"

```

```

FEATURES
source
1..699
/organism="Homo sapiens"

```

```

BASE COUNT      135 a      270 c      189 g      104 t      1 others
ORIGIN
Alignment Scores:
Pred. No.:      1,94e-86      Length:      699
Score:      1097.00      Matches:      209
Percent Similarity:      100.00%      Conservative:      0
Best Local Similarity:      100.00%      Mismatches:      0
Query Match:      63.37%      Indels:      0
DB:      12      Gaps:      0
US-09-709-103-3f1 (1-332) x BG706012 (1-699)
Qy      3 ProSerGlyAlaGlyAlaProSerProSerArgAlaGlnProGlnGlnSerProProAla 22
Db      71 CCGAGCGAGAGCCGAGACCCCAAGCCGAGCCCGCCGACCCGAGCAGAGAGCCCTCAGC 130
Qy      23 AlaH1sProAlaGlyH1sProSerAspProGlnProLeuSerAlaLeuLeuSerAlaPro 42
Db      131 GCTCACCCCGGTGACACCCACGACCTTCAGCCGCTCTGCTGCTTCGCGCCG 190
Qy      43 ArgProProSerArgProLeuCysPrometLysLeuAlaAlaMetLysLysMetCys 62
Db      191 CGCCCGCCCTCGGGGCCCTCTGCCCCAATGAACCTGGCCGATGATCAAGAAAGATGC 250
Qy      63 ProSeraspSerGlnLeuSer11leProAlaLysasnCysTyArgMetVal11leuGly 82
Db      251 CCGAGCAGCTCGAGGCTGAGTATCCCGGCCAAGACTGCTATGCAATGATCTCTCGG 310
Qy      83 SerSerLysValGlyLysThrAla11leValSerArgPheLeuThrGlyArgPheGluasp 102
Db      311 TCGTCAAGGTGGCAAGAGCGCCATGCTGCGGCTCTTCACCGCGCTTCGAGAGC 370
Qy      103 AlaTyThrProthrilleguaspPheH1sArgLysPheTySer11leargLylgluVal 122
Db      371 GCTTACAGCGCTTACATCGAGGACTTCCACCGCAAGTCTTACATCCGCGCGAGGTG 430
Qy      123 TyrGlnleuasp11leuaspThrSerGlyAsnH1sProPheProAlaMetArgArgLeu 142
Db      431 TACAGCTCGACATCTCGACACTCCGCAACCAACCTTCCCGGCAAGCGCGCTC 490
Qy      143 Ser11leuThrGlyaspValPhe1leuValPheSerLeuaspAsnArgAspSerPhe 162
Db      491 TCATCTCCACAGAGAGAGCTTTCATCTGTTGATCTGAGTCTGCAACCCGACTCTTC 550
Qy      163 GlnGluValGlnArgLeuArgGlnGln11leuaspThrLysSerCysleuLysasnLys 182
Db      551 GAGGAGGTGCGAGCGCTCAGACAGATCTTCACACCAAGTCTTCCCTCAAGAACAA 610
Qy      183 ThrLysGluaspValaspValProleuVal11leCysglYasnLysglYaspArgAspPhe 202
Db      611 ACCAAGAGAACGTGAGAGTCCCTGCTGATCTGCGGCAACAAAGGTGACCGGACTTC 670
Qy      203 TyArgGluValaspGlnArgGlu11le 211
Db      671 TACCGGAGGTGACCAAGCGGAGATC 697
RESULT 15
BG709229

```

LOCUS BG709229 699 bp mRNA linear EST 07-MAY-2001
DEFINITION 602674615F1 NIH_MGC_96 Homo sapiens CDNA clone IMAGE:4797482 5',
mRNA sequence.
ACCESSION BG709229
VERSION BG709229.1 GI:13987357
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 699)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHAM10683 row: m column: 03
High quality sequence stop: 699.
FEATURES
source location/Qualifiers
1..699
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4797482"
/clone_lib="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/notes="Organ: brain; Vector: pBluescriptR (modified
pBluescript KS+); Site 1: BamHI; Site 2: SalI-xhoI (gtcagag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',
size-selected for average insert size 2.3 kb and
normalized to ROT 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT 136 a 270 c 188 g 105 t
ORIGIN

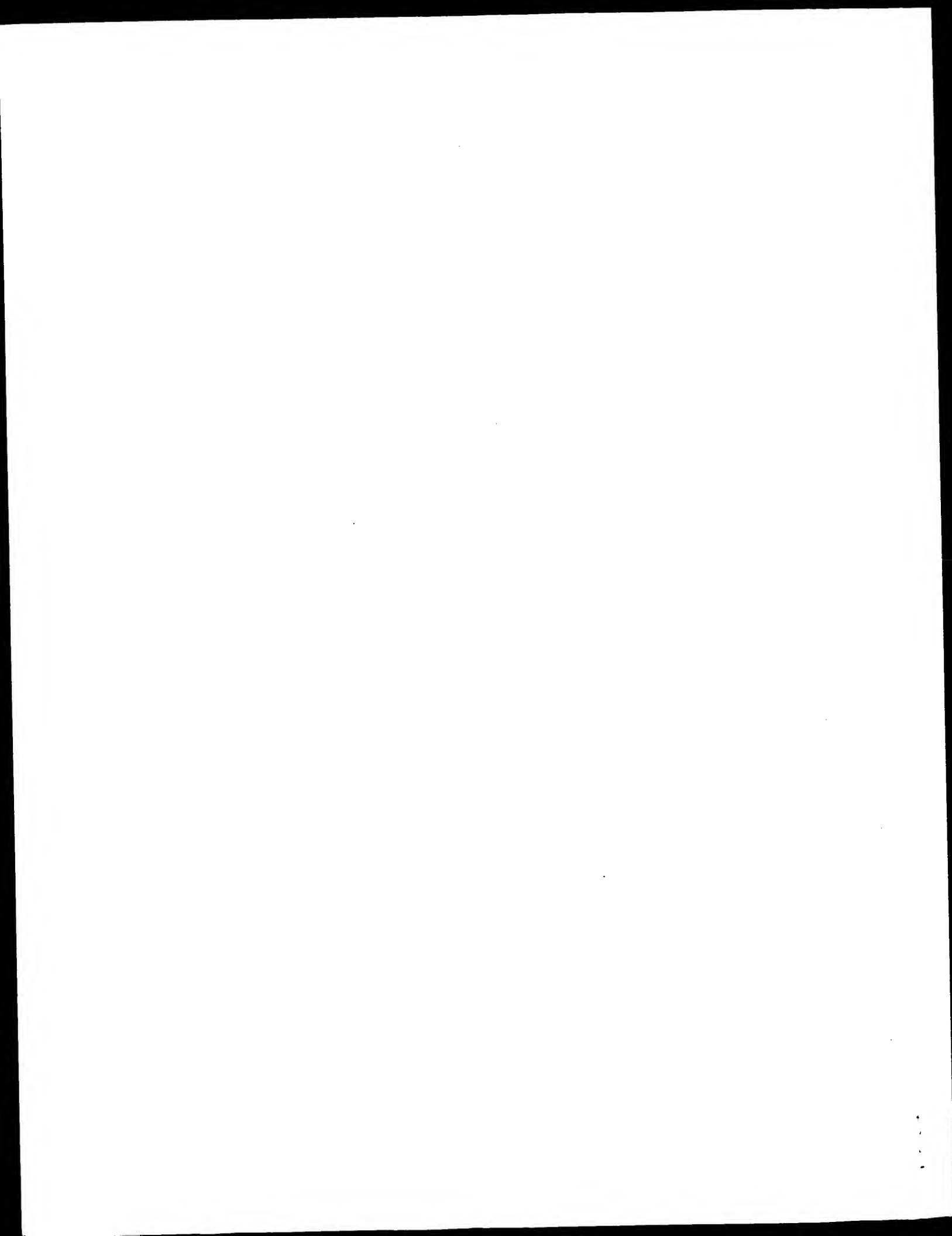
Alignment Scores:
Pred. No.: 3.71e-84 Length: 699
Score: 1071.00 Matches: 209
Percent Similarity: 99.05% Conservative: 0
Best Local Similarity: 99.05% Mismatches: 0
Query Match: 61.87% Indels: 2
DB: 12 Gaps: 0

US-09-709-103-3F1 (1-332) x BG709229 (1-699)

QY 3 ProSerGlyAlaGlyAlaProSerProSerArgAlaGln-ProGluGlnSerProAl 22
Db 71 CCGAGCGGAGCGGAGCGCCCAAGCCGAGCCGCCGCCAGCCCTGAGCAGCCCTCCAGC 130
QY 22 aAlaHisProAlaCysHisProSerAspProGlnProLeuSerAlaLeuSerAlaPr 42
Db 131 CGCTACCCCGCGGTGCCACCCAGGAGCCCTCAGCCGCTCTCTGCCCTCTCTCGGCCCC 190
QY 42 oArgProProSerArgProLeuCysProMetIysLeuAlaAlaMetIleIysLysMetCy 62
Db 191 GCGCCCGCCCTCGCGGCCCTCTGCCCAATGAACTGGCCGCGATGATCAAGAAGATGTG 250
QY 62 sProSerAspSerGluLeuSerIleProAlaIysAsnCysTyrArgMetValIleLeuGl 82
Db 251 CCGGAGCGACTCGGAGCTGAGTATCCCGGCCAAGACTGTATCGCATGGTATCTCTCGG 310
QY 82 y-SerSerIysValGlyLysThrAlaIleValSerArgPheLeuThrGlyArgPheGluA 102

Db 311 ACTCGTCCAAAGTGGGGAAGAGCGCCATCGTGTGCGCTTCTCACCGGCGGCTTCGAGG 370
QY 102 spAlaTyrThrProThrIleGluAspPheHisArgLysPheTyrSerIleArgGlyGluV 122
Db 371 ACGCCTACAGCGCTACCATCGAGGACTTCCACCGCAAGTTCTACTCTCATCCGCGGCGAGG 430
QY 122 alTyrGlnLeuAspIleLeuAspThrSerGlyAsnHisProPheProAlaMetArgArgL 142
Db 431 TCTACCACTCGACATCTCTGACAGCTCCGCGCAACACCGCTTCCCGCCATCGCGGCC 490
QY 142 euSerIleLeuThrGlyAspValPheIleLeuValPheSerLeuAspAsnArgAspSerP 162
Db 491 TCTCCATCTCACAGGAGACGTTTTCATCTGCTGTTTTCAGTCTGGACCAACCGGACTCCT 550
QY 162 heGluGluValGlnArgLeuArgGlnGlnIleLeuAspThrLysSerCysLeuLysAsnL 182
Db 551 TCGAGGAGTGCAGCGGCTCAGGCGAGCAGATCTCTCGACACCAAGTCTTGCTCAAGAACA 610
QY 182 ysThrLysGluAsnValAspValProLeuValIleCysGlyAsnLysGlyAspArgAsp 202
Db 611 AAACCAAGGAGAACCTGGACGTCGCCCTGCTGTCATCTGCGGCAACAAAGGTGACCGGACT 670
QY 202 heTyrArgGluValAspGlnArgGluIle 211
Db 671 TCTACCGGAGGTGGACCGCGGAGATC 699

Search completed: December 30, 2002, 19:42:40
Job time : 2189 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 30, 2002, 18:06:06 ; Search time 76 Seconds
(without alignments)
1339.693 Million cell updates/sec

Title: US-09-709-103-3F1
Perfect score: 1731
Sequence: 1 GIPSCAGAPSPRAQPEOSP.....IREKASAGSQAQKRCRVS 332

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-O=/cgn2_1/USPNO_spool/SULLIVAN3F1/runat_30122002.144155.22165/app_query.fasta_1.519
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCI=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=SULLIVAN3F1@cgn 1.1.31 @runat_30122002.144155.22165 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=10 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA:
1: /cgn2_6/prodata/1/ina/5A.COMB.seq:
2: /cgn2_6/prodata/1/ina/5B.COMB.seq:
3: /cgn2_6/prodata/1/ina/6A.COMB.seq:
4: /cgn2_6/prodata/1/ina/6B.COMB.seq:
5: /cgn2_6/prodata/1/ina/PTUS.COMB.seq:
6: /cgn2_6/prodata/1/ina/backfiles1.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1713	99.0	1841	4	US-09-053-374A-1
2	1651	95.4	3986	4	US-09-053-374A-3
3	1435.5	82.9	1689	4	US-09-053-374A-4
4	1388.5	80.2	3079	4	US-09-053-374A-6
5	300	17.3	615	1	US-08-247-946A-5
6	300	17.3	615	5	PCT-US95-06420-5
7	279	16.1	5775	1	US-08-306-691B-15
8	279	16.1	5775	5	PCT-US93-06251-29
9	277.5	16.0	570	4	US-08-884-866A-2
10	277.5	16.0	570	4	US-08-884-866A-11
11	276	15.9	607	2	US-08-429-964-85
12	275.5	15.9	480	4	US-08-884-866A-9
					Sequence 1, Appli
					Sequence 3, Appli
					Sequence 4, Appli
					Sequence 5, Appli
					Sequence 6, Appli
					Sequence 7, Appli
					Sequence 8, Appli
					Sequence 9, Appli
					Sequence 10, Appli
					Sequence 11, Appli
					Sequence 12, Appli
					Sequence 13, Appli
					Sequence 14, Appli
					Sequence 15, Appli
					Sequence 16, Appli
					Sequence 17, Appli
					Sequence 18, Appli
					Sequence 19, Appli
					Sequence 20, Appli
					Sequence 21, Appli
					Sequence 22, Appli
					Sequence 23, Appli
					Sequence 24, Appli
					Sequence 25, Appli
					Sequence 26, Appli
					Sequence 27, Appli
					Sequence 28, Appli
					Sequence 29, Appli
					Sequence 30, Appli
					Sequence 31, Appli
					Sequence 32, Appli
					Sequence 33, Appli
					Sequence 34, Appli
					Sequence 35, Appli
					Sequence 36, Appli
					Sequence 37, Appli
					Sequence 38, Appli
					Sequence 39, Appli
					Sequence 40, Appli
					Sequence 41, Appli
					Sequence 42, Appli
					Sequence 43, Appli
					Sequence 44, Appli
					Sequence 45, Appli

13	275.5	15.9	4480	4	US-09-167-322-12	Sequence 12, Appli
14	266.5	15.4	450	4	US-08-884-866A-10	Sequence 10, Appli
15	266.5	15.4	574	2	US-08-429-964-83	Sequence 83, Appli
16	265.5	15.3	2436	1	US-08-306-691B-16	Sequence 16, Appli
17	237	13.7	2309	3	US-09-078-317-3	Sequence 3, Appli
18	237	13.7	2309	3	US-09-454-818-3	Sequence 3, Appli
19	234.5	13.5	600	3	US-09-078-317-1	Sequence 1, Appli
20	234.5	13.5	600	3	US-09-454-818-1	Sequence 1, Appli
21	220	12.7	3497	4	US-09-503-505A-2	Sequence 2, Appli
22	219.5	12.7	563	4	US-09-385-982-426	Sequence 426, App
23	203	11.7	897	2	US-09-006-535-2	Sequence 2, Appli
24	203	11.7	1525	2	US-09-006-535-1	Sequence 1, Appli
25	201	11.6	6453	1	US-08-306-691B-14	Sequence 14, Appli
26	201	11.6	6453	3	US-09-209-668-10	Sequence 10, Appli
27	201	11.6	6453	3	US-09-356-952-8	Sequence 8, Appli
28	198	11.4	897	2	US-09-006-535-7	Sequence 7, Appli
29	193	11.1	1443	1	US-08-076-089-1	Sequence 1, Appli
30	193	11.1	1443	2	US-08-707-200-1	Sequence 1, Appli
31	193	11.1	1443	4	US-08-896-565-1	Sequence 1, Appli
32	193	11.1	1443	5	PCT-US93-05643-1	Sequence 1, Appli
33	188	10.9	702	3	US-08-842-976-2	Sequence 2, Appli
34	188	10.9	702	3	US-09-213-397-2	Sequence 2, Appli
35	188	10.9	702	3	US-09-416-489-2	Sequence 2, Appli
36	187	10.8	1098	2	US-08-948-616-6	Sequence 6, Appli
37	187	10.8	1098	2	US-09-193-510-6	Sequence 6, Appli
38	187	10.8	1098	4	US-09-368-402-6	Sequence 6, Appli
39	185.5	10.7	2964	2	US-08-846-790A-2	Sequence 2, Appli
40	185.5	10.7	2964	3	US-08-935-333-2	Sequence 2, Appli
41	178	10.3	1074	2	US-09-156-424-1	Sequence 1, Appli
42	178	10.3	1074	4	US-09-387-341-1	Sequence 1, Appli
43	177.5	10.3	1166	5	PCT-US96-12129B-1	Sequence 1, Appli
44	175.5	10.1	603	4	US-09-325-932A-29	Sequence 29, Appli
45	175.5	10.1	932	4	US-09-325-932A-28	Sequence 28, Appli

ALIGNMENTS

RESULT 1

US-09-053-374A-1
; Sequence 1, Application US/09053374A
; Patent No. 6462177
; GENERAL INFORMATION:
; APPLICANT: YEN, KWANG-MU
; TITLE OF INVENTION: MAMMALIAN BLOOD LOSS-INDUCED GENE, KD312
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: AMGEN INC.
; STREET: ONE AMGEN CENTER DRIVE
; CITY: THOUSAND OAKS
; STATE: CA
; COUNTRY: US
; ZIP: 91320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/053,374A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: COOK, ROBERT R.
; REGISTRATION NUMBER: 31,602
; REFERENCE/DOCKET NUMBER: A-514
; INFORMATION FOR SEQ ID NO. 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1841 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:

NAME/KEY: CDS
LOCATION: 255..1097
US-09-053-374A-1

Alignment Scores:
Pred. No.: 1,156-146 Length: 1841
Score: 1713.00 Matches: 329
Percent Similarity: 99.70% Conservative: 0
Best Local Similarity: 98.96% Mismatches: 1
Query Match: 4 Indels: 0
Gaps: 0

US-09-709-103-3f1 (1-332) x US-09-053-374A-1 (1-1841)

QY 3 ProSerGlyAlaGlyAlaProSerProSerArgAlaGlnProGluGlnSerProProAla 22
DB 108 CCGAGCGAAGCCGGAGCCCAAGCCCGAGCCGCCCGAGCAGAGCCCTCCAGCC 167
QY 23 AlaHisProAlaCyHisProSerAspProGlnProLeuSerAlaLeuLeuSerAlaPro 42
DB 168 GCTCAACCCGCGTGGCCACCCCGAGCAGCCCTCCAGCCCTCTCTGCCCCCTCTCGGCCCG 227
QY 43 ArgProProSerArgProLeuCyProMetLysLeuAlaMetLysLysMetCys 62
DB 228 CGCCCGCCCTCGCGCCCTCTCCCAATGAAGCTGGCCGATGATCAAGAGATGTC 287
QY 63 ProSerAspSerGluLeuSerLleProAlaLysAsnCyTyArgMetValLleuGly 82
DB 288 CCGAGGCACTCGAGGCTAGTATCCCGCCAGAACTGCTATCGATGTCATCTCTCGGC 347
QY 83 SerSerLysValGlyLysThrAlaIleValSerArgPheLeuThrglyArgPheGluAsp 102
DB 348 TCGTCAGAGTGGCCAGAGCGCATCTGTCTGCTTCCTCAACCGCGCTTCGAGAGC 407
QY 103 AlaTyThrProThrLleGluAspPheHisArgLysPheTySerLleArgGlyGluVal 122
DB 408 GCCACACGCTTACCTACGAGACTTCCACCGCAAGTTCATCTCCATCCGCGGAGGTC 467
QY 123 TyGlnLeuAspLleLeuAspThrSerGlyAsnHisProPheProAlaMetArgLysLeu 142
DB 468 TACCAAGCTCGACATCTCGACATCTCCGACACCAACCGCTCCCGCATGCGGCCCTTC 527
QY 143 SerLleLeuThrGlyAspValPheLleLeuValPheSerLleAspAsnArgAspSerPhe 162
DB 528 TCATCTCTCAACAGAGAGATTTTCATCTCTGCTTCAGTCTGACCAACCGCATCTCTTC 587
QY 163 GluGluValGlnArgLysGlnGlnLleLeuAspThrLysSerCysLeuLysAsnLys 182
DB 588 GAGGAGGTGCGACCGCTCAGGAGCAGATCTCTCGACACCAAGTCTTCCCTCAAGAACAA 647
QY 183 ThrLysGluAsnValAspValProLeuValLleCysGlyAsnLysGlyAspArgAspPhe 202
DB 648 ACCAAGAGAGATGTGAGCGTCCCTGCTCATCTGCGGCAACAGGGGTGAGCGGACTTC 707
QY 203 TyArgGluValAspGlnArgGlnLleGluGlnLeuValGlyAspAspProGlnArgCys 222
DB 708 TACCGGAGGTGAGCAGCGGCTCAGGAGCAGATCTCTCGACACCAAGTCTTCCCTCAAGAACAA 767
QY 223 AlaTyThrPheGluLleSerAlaLysLysAsnSerSerLleAspGlnMetPheArgAlaLeu 242
DB 768 GCCTACTTGAATCTCGGCCAAGAACACACAGCTGAGACCAAGTCTTCCGCGCTC 827
QY 243 PheAlaMetAlaLysLeuProSerGluMetSerProAspLeuHisArgLysValSerVal 262
DB 828 TTGCGCATGGCCCAAGCTGCGCCAGCGAGATGAGCCCGACCTCAACCGCAAGTCTCGGTG 887
QY 263 GlnTyTyCysAspValLeuHisLysLysAlaLeuArgAsnLysLysLeuLeuArgAlaGly 282
DB 888 CAGTACGTCACGCTGTCGCAAGAGAGCCCTCGGAAACAGAGCTGCTGCGGCGCGC 947
QY 283 SerGlyGlyGlyGlyAspProGlyAspAlaPheGlyLleValAlaProPheAlaArg 302
DB 948 AGCGGCGGCGCGCGCGAGACCGGCGGAGCGCTTGGCATCGTGGACCTTTCGCGGC 1007

QY 303 ArgProSerValHisSerAspLeuMetTyrlleArgGlyLysAlaSerAlaGlySerGln 322
DB 1008 CGGCCGAGCTGACACAGCAGCCTCATGTATCATCTCCGAGAGGCCAGCGCGGAGCCAG 1067
QY 323 AlaLysAspLysGluArgCysValLleSer 332
DB 1068 GCCAAGAGCAAGAGCGCTGCTCATCAGC 1097

RESULT 2
US-09-053-374A-3
Sequence 3, Application US/09053374A

Patent No. 6462177
GENERAL INFORMATION:
APPLICANT: YEN, KWANG-MU
TITLE OF INVENTION: MAMMALIAN BLOOD LOSS-INDUCED GENE, KD312
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: AMGEN INC.
STREET: ONE AMGEN CENTER DRIVE
CITY: THOUSAND OAKS
STATE: CA
COUNTRY: US
ZIP: 91320
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/053,374A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: COOK, ROBERT R.
REGISTRATION NUMBER: 31,602
REFERENCE/DOCKET NUMBER: A-514
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3986 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-09-053-374A-3

Alignment Scores:

Pred. No.: 1,496-140 Length: 3986
Score: 1651.00 Matches: 329
Percent Similarity: 82.04% Conservative: 0
Best Local Similarity: 95.38% Mismatches: 1
Query Match: 4 Indels: 71
Gaps: 1

US-09-709-103-3f1 (1-332) x US-09-053-374A-3 (1-3986)

QY 3 ProSerGlyAlaGlyAlaProSerProSerArgAlaGlnProGluGlnSerProProAla 22
DB 627 CCGAGCGAAGCCGGAGCCCAAGCCCGAGCCGCCCGAGCAGAGCCCTCCAGCC 686
QY 23 AlaHisProAlaCyHisProSerAspProGlnProLeuSerAlaLeuLeuSerAlaPro 42
DB 687 GCTCAACCCGCGTGGCCACCCCGAGCAGCCCTCCAGCCCTCTCTGCCCCCTCTCGGCCCG 746
QY 43 ArgProProSerArgProLeuCyProMetLysLeuAlaMetLysLysMetCys 62
DB 747 CGCCCGCCCTCGCGCCCTCTCCCAATGAAGCTGGCCGAGATGATCAAGAGATGTC 806
QY 63 ProSerAspSerGluLeuSerLleProAlaLysAsnCyTyArgMetValLleuGly 82
DB 807 CCGAGGCACTCGAGGCTAGTATCCCGGCAAGAACTGCTATCCATGCTATCTCTCGGC 866
QY 83 SerSerLysValGlyLysThrAlaIleValSerArgPheLeuThrglyArgPheGluAsp 102

Db	867	TCGTC	CAAGTGGGCAAGCGGCATCGTGTGCGCTTCTCA	CGCGCGCTTCAGAGAC	926
Qy	103	AlaTyrThr	ProThrIleGluAspPheHisArgIysPheTyrSerIle	ArgGlyVal	122
Db	927	GCCTAC	AGCGCTACCATCGAGGACTTCACCGCAAGTCTACTCTCATCGGCGGAGGTC	986	
Qy	123	TyrGlnLeu	AspIleLeuAspThrSerGlyAsnHisProPheProAlaMetArgArgLeu	142	
Db	987	TACCACTCGACATCTCGACACGTCGGCAACCA	CCCGTTCCCGGCCATCGCGGCGCTC	1046	
Qy	143	SerIleLeuThr	-----	-----	146
Db	1047	TCCATCTCTACAGGTACGCCGGGGCCGGCAGGTG	CGGGAGGAAAGGCGGGCAACCTC	1106	
Qy	146	-----	-----	-----	146
Db	1107	CGGCCAGGCGCCCGCGAGCGCGTCCGGTGTCTCGGCGTGTGCGCGCCAGTAGTCGGCTTCGCG	1166		
Qy	146	-----	-----	-----	146
Db	1167	CTTAGAGAGGTAGCGCGCCCGCGCGCTCAAAGTCAGCCGACTTGTCCCTGGGCG	1226		
Qy	147	-----	-----	-----	146
Db	1227	GCCACCTCACCTTCTCTCTTTCTGCTCTGTGTCGCCCTCTAGGAGACGTTTCATCCT	1286		
Qy	152	uValPheSerLeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGlnGlnIle	172		
Db	1287	GGTGTTCAGTCTGGACAACCGCGACTCTCTCGAGGAGGTGCAGCGCTCAGGCGAGCAT	1346		
Qy	172	eLeuAspThrIysSerCysLeuLysAsnLysThrLysGluValAsnValAspValProLeuVal	192		
Db	1347	CCTCGACACCAAGCTTGCTCTALAGAACAAACCAAGAGAACGTGCAGCTGCCCTGT	1406		
Qy	192	IleCysGlyAsnLysGlyAspArgAspPheTyrArgGluValAspGlnArgGluIleG	212		
Db	1407	CATCTGCGGCACAAGGTGACCGGACTTCTACCGGAGTGGACCGCGAGATCGA	1466		
Qy	212	uGlnLeuValGlyAspAspProGlnArgCysAlaTyrPheGluIleSerAlaLysLysAs	232		
Db	1467	CGAGCTGTGGCGGACGACCCCGCGCTCGCTACTTCAGATCTCGGCCCAAGAA	1526		
Qy	232	nSerSerLeuAspGlnMetPheArgAlaLeuPheAlaMetAlaLysLeuProSerGluMe	252		
Db	1527	CAGCAGCTGGACCAGATGTTCCGGCGCTCTTCGCCATGGCCACAGCTCCCGAGAGAT	1586		
Qy	252	tSerProAspIleuHisArgLysValSerValGlnTyrCysAspValLeuHisLysLysAl	272		
Db	1587	GAGCCACAGACTGCACCGCAAGGTCTCGGTGAGTACTGCGACGTGCTGCACAAGAGGC	1646		
Qy	272	aLeuArgAsnLysLysLeuLeuArgAlaGlySerGlyGlyGlyGlyAspProGlyAs	292		
Db	1647	GCTGCGAAACAAGAGCTGCTGCGGCGGCGAGCGGCGCGCGGCGACCCGCGGA	1706		
Qy	292	pAlaPheGlyIleValAlaProPheAlaArgProSerValHisSerAspLeuMetTy	312		
Db	1707	CGCCTTTGGCATCGTGGCACCTTTCGCGCGCGGCCCGCGGTACACAGGACCTCATGTA	1766		
Qy	312	rIleArgGluLysAlaSerAlaGlySerGlnAlaLysAspLysGluArgCysValIleSe	332		
Db	1767	CATCCGGAGAGGCCGCGCGGACCGAGCCAGGCCAAGGACAAGGAGCGCTGCGTCATCAG	1826		
Qy	332	r	332		
Db	1827	C	1827		

RESULT 3

RESULT 3
US-09-053-374A-4

00 00 000-374A-4
; Sequence 4, Application US/09053374A

; Patent No. 6462177

; GENERAL INFORMATION:

APPLICANT: YEN, KWANG-MU

Db	420	GACGTTTTCATCTCGGGTTCAGACTTAGCAACCCGGACTCTCTTGAGAGGTGCAGAAAGG	479
QY	168	LeuArgGlnGlnIleLeuAspThrIlySerCysLeuIlyAsnIlyThrIlySGluAsnVal	187
Db	480	CTCAAAACGACGAGATCCCTTAGACACCAAGATCTGTCTCAAGAACAAACCAAGAGATGG	539
QY	188	AspValProLeuValIleCysGlyAsnIlySGIlyAspArgAspPheIlyArgGluValAsp	207
Db	540	GACCTGGCCGCTGGTCATTTTCGGTAAACAAAGGGACCCGGGACTTCACCCGCAAGTGGAG	599
QY	208	GlnArgGlnIleGlnGlnLeuValGlyAspAspProGlnIlyArgValIlyIlyPheGlnIle	227
Db	600	CACGGGGAGATTGAGCAGCTGTGGCGATGACCTCAGGGTGTCTTACCTTCAGATC	659
QY	228	SerAlaIlySlyAsnSerSerLeuAspGlnMetPheArgIleLeuPheAlaMetAlaIlyS	247
Db	660	TCGGCCAGAGAAATAGCAGGCTTGACACAGATGTTCCGTCCGCTCTTTGGCAGTGGCAAG	719
QY	248	LeuProSerGlnMetSerProAspLeuHISArgIlyValSerValGlnIlyCysAspVal	267
Db	720	CTGCCTTACCGAGATGAGCCCTTGACTTGACACCGCAAGGTGTCTGGACGATCTGTGACGTG	779
QY	268	LeuHISlySlyAlaLeuLeuArgAsnIlySlyLeuLeuArgAlaGlySerGlyGlyGlyGly	287
Db	780	CTGCACAAAAAGGCTCTGAGGAAACAAGAACTTCGCTGGCTGGGGCACAC---GAGGTGGG	836
QY	288	GlyAspProGlyAspAlaPheGlyIleValAlaProPheAlaArgArgProSerValHis	307
Db	837	GGCGACCAAGGAGATGCTTGGCACTCTGGCGCCCTTGTCTCCGACACTAAGGTGCAT	896
QY	308	SerAspLeuMetIlyIleArgGlnIlySlyAsnSerAlaGlySerGlnAlaIlyAspIlyGln	327
Db	897	AGGACCTCATGATCACTTGTGAGAAACAGGTGCACACCAAGGCTTAAGACAAAGAG	956
QY	328	ArgCysValIleSer332	
Db	957	CGCTGTCTATCACT971	

RESULT 4
 US-09-053-374A-6
 : Sequence 6, Application US/09053374A
 : Patent No. 6462177
 :
 : GENERAL INFORMATION:
 : APPLICANT: YEN, KWANG-MU
 : TITLE OF INVENTION: MAMMALIAN BLOOD LOSS-INDUCED GENE, KD312
 :
 : NUMBER OF SEQUENCES: 9
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: AMGEN INC.
 : STREET: ONE AMGEN CENTER DRIVE
 : CITY: THOUSAND OAKS
 : STATE: CA
 : COUNTRY: US
 : ZIP: 91320
 :
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: Patent In Release #1.0, Version #1.30
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/09/053,374A
 : FILING DATE:
 : CLASSIFICATION: 435
 : ATTORNEY/AGENT INFORMATION:
 : NAME: COOK, ROBERT R.
 : REGISTRATION NUMBER: 31,602
 : REFERENCE/DOCKET NUMBER: A-514
 : INFORMATION FOR SEQ ID NO: 6:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 3079 base pairs
 : TYPE: nucleic acid
 : STRANDEDNESS: single
 : TOPOLOGY: linear

[illegible]

Qy 258 ArgLysValSerValGlnTyrCysAspValLeuHisLysLysAlaLeuArgAsnLysLys 277
 Db 1058 CGCAAGGTGTCTGTGACAGTACTGTGACGCTCTGCACAAAAGGCTCTGAGGAACAGAG 1117
 Qy 278 LeuLeuArgAlaGlySerGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 297
 Db 1118 CTTCTGCTGGCGGACGAC---GGAGGTGGGGGGGACACGAGAGTGCCTTTGGCATCTTG 1174
 Qy 298 AlaProPheAlaArgArgProSerValHisSerAspLeuMetTyrIleArgGluLysAla 317
 Db 1175 GCGCCCTTTGCTCGCAGACCTACGCTGCATAGCAGCCTCATGTACATTCTGTGAGAAACC 1234
 Qy 318 SerAlaGlySerGlnAlaLysAspLysGluArgCysValIleSer 332
 Db 1235 AGGTGTACGACGACCGAGGTAAAGCAAGGAGCGCTGTGTATCATCAGT 1279

RESULT 5

US-08-247-946A-5
 ; Sequence 5, Application US/08247946A
 ; Patent No. 5792638
 ; GENERAL INFORMATION:
 ; APPLICANT: AARONSON, S.A.; CHAN, A.;
 ; APPLICANT: MIKI, T.
 ; TITLE OF INVENTION: NOVEL HUMAN RAS-RELATED
 ; TITLE OF INVENTION: ONCOGENES UNMASKED BY EXPRESSION OF CDNA
 ; TITLE OF INVENTION: CLONING
 ; NUMBER OF SEQUENCES: 11
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: MORGAN & FINNEGAN
 ; STREET: 345 PARK AVENUE
 ; CITY: NEW YORK
 ; STATE: NEW YORK
 ; COUNTRY: USA
 ; ZIP: 10154

COMPUTER READABLE FORM:
 MEDIUM TYPE: FLOPPY DISK
 COMPUTER: IBM PC COMPATIBLE
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: WORDPERFECT 5.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/247,946A
 FILING DATE: 24-MAY-1994
 CLASSIFICATION: 536
 ATTORNEY/AGENT INFORMATION:
 NAME: DOROTHY R. AUTH
 REGISTRATION NUMBER: 36,434
 REFERENCE/DOCKET NUMBER: 2026-4150
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 758-4800
 TELEFAX: (212) 751-6849
 TELEX: 421792
 INFORMATION FOR SEQ ID NO: 5:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 615
 TYPE: Nucleic acid
 STRANDEDNESS: Double
 TOPOLOGY: Unknown
 MOLECULE TYPE: CDNA
 HYPOTHETICAL: No
 ORIGINAL SOURCE:
 ORGANISM: Human
 STRAIN:
 INDIVIDUAL ISOLATE:
 DEVELOPMENTAL STAGE:
 HAPLOTYPE:
 TISSUE TYPE:
 CELL TYPE:
 CELL LINE:
 ORGANELLE:
 FEATURE:
 NAME/KEY: TC21 gene
 LOCATION:
 IDENTIFICATION METHOD:

OTHER INFORMATION:

US-08-247-946A-5

Alignment Scores:

Pred. No.: 8,278-19 Length: 615
 Score: 300.00 Matches: 71
 Percent Similarity: 55.32% Conservative: 33
 Best Local Similarity: 37.77% Mismatches: 70
 Query Match: 17.33% Indels: 14
 DB: 1 Gaps: 3

US-09-709-103-3f1 (1-332) x US-08-247-946A-5 (1-615)

Qy 76 TyrArgMetValIleLeuGlySerSerLysValGlyLysThrAlaIleValSerArgPhe 95
 Db 43 TACCGGCTCGTGTGTGGCGGGCGGCGTGGCAAGTCGGCGCTCACCATCAGTTTC 102
 Qy 96 LeuThrGlyArgPheGluAspAlaTyrThrProThrIleGluAspPheHisArgLysPhe 115
 Db 103 ATCCAGTCTCTATTTTGTAAACGGATTATGATCCAAACCATTTGAAGATTCTTACACAAAGCAG 162
 Qy 116 TyrSerIleArgGlyGluValTyrGlnLeuAspIleLeuAspThrSerGlyAsnHisPro 135
 Db 163 TGTGTGATAGATGACAGACAGCCCGCTAGATATTTGGATACAGCAGGANNNGAAGAG 222
 Qy 136 PheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIleLeuValPheSer 155
 Db 223 TTTGGAGCCATGAGAGAACAGTATATGAGGACTGCGCAAGCTTCTGTGGTCTTTTCA 282
 Qy 156 LeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGlnGlnIleLeuAspThr 175
 Db 283 GTCACAGATAGAGCAGTTTGAAGAAATCTATAAGTTTCAAAGACAGATTTCTC 336
 Qy 176 LysSerCysLeuLysAsnLysThrLysGluAsnValAspValProLeuValIleCysGly 195
 Db 337 -----AGAGTAAAGGATCGTGATGAGTTCCCAATGATTTTAATTGGT 378
 Qy 196 AsnLysGlyAspArgAspPheTyrArgGluValAspGlnArgGluIleGluGlnLeuVal 215
 Db 379 AATAAAGCAGATCTGGATCATCAAAAGACAGGTAAACACAGGAAGAGGACACAGTTAGCA 438
 Qy 216 GlyAspAspProGlnArgCysAlaTyrPheGluIleSerAlaLysLysAsnSerSerLeu 235
 Db 439 CGGCAG-----CTTAAGGTAACATACATCGAGGCGATCAGCAAAAGATTAGGATGAATGA 492
 Qy 236 AspGlnMetPheArgAlaLeuPheAlaMetAlaLysLeuProSerGluMet----- 252
 Db 493 GATCAAGCTTTCCATGAACCTGTCCGGGTTATCAGGAAATTTCAAGAGCAGGAATGCTCCT 552
 Qy 253 ---SerProAspLeuHisArgLys 259
 Db 553 CTTTACCACGAACCAACACGGA 576

RESULT 6

PCT-US95-06420-5
 ; Sequence 5, Application PC/TUS9506420
 ; GENERAL INFORMATION:
 ; APPLICANT: AARONSON, S.A.; CHAN, A.;
 ; APPLICANT: MIKI, T.
 ; TITLE OF INVENTION: NOVEL HUMAN RAS-RELATED
 ; TITLE OF INVENTION: ONCOGENES UNMASKED BY EXPRESSION OF CDNA
 ; TITLE OF INVENTION: CLONING
 ; NUMBER OF SEQUENCES: 5
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: MORGAN & FINNEGAN
 ; STREET: 345 PARK AVENUE
 ; CITY: NEW YORK
 ; STATE: NEW YORK
 ; COUNTRY: USA
 ; ZIP: 10154
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: FLOPPY DISK
 ; COMPUTER: IBM PC COMPATIBLE

```

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/06420
FILING DATE: 24-MAY-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/247,946
FILING DATE: 24-MAY-1994
ATTORNEY/AGENT INFORMATION:
NAME: DOROTHY R. AUTH
REGISTRATION NUMBER: 36,434
REFERENCE/DOCKET NUMBER: 2026-4150PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 615
TYPE: Nucleic acid
STRANDEDNESS: Double
TOPOLOGY: Unknown
MOLECULE TYPE: cDNA
HYPOTHETICAL: No
ORIGINAL SOURCE:
ORGANISM: Human
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE:
CELL LINE:
ORGANELLE:
FEATURE:
NAME/KEY: TC21 gene
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION:
PCT-US95-06420-5

Alignment Scores:
Pred. No.: 8,276-19 Length: 615
Score: 300.00 Matches: 71
Percent Similarity: 55.32% Conservative: 33
Best Local Similarity: 37.77% Mismatches: 70
Query Match: 17.33% Indels: 14
DB: 5 Gaps: 3

US-09-709-103-3f1 (1-332) x PCT-US95-06420-5 (1-615)
QY 76 TYRARGMETVALLELENGLYSERISERLYSVALGLYVSTHRAALALEVALSERARGPHE 95
DB 43 TACCGGCTCGTGTGTGGCGGGCGGGCGGCGGCAATCGCGCGGTCAACATTCAGTTTC 102
QY 96 LEUETHGILYARGPHEGLUASPALATYRTHRPROTHIRLEGLUASPHEHISARGLYSPHE 115
DB 103 ATCCAGTCCATTGTTGAACGATTAATCCCAACATTCGAAAGATTCTTAACAAGACG 162
QY 116 TYRSEITLHARGGLYGLUVALTYRGLINLEUASPHEISERGILYASNHISPRO 135
DB 163 TGTGTGATATGATGACAGAGCAGCCCGCTAGATATTGATGATGACAGCAGANNNGAAG 222
QY 136 PHEPROALHAEIRGARGLEUSERILELEUTHRGILYASPVALPHEILEUVALPHESER 155
DB 223 TTGGAGCCCTAGACAGAACAGTATATAGAGCTGCGCAAGCTTCCTGTTGTTTCA 282
QY 156 LEUASPANHRGASPSERPIHEGLUVALGLNARGLEUARGINGINILEUASPTHR 175
DB 283 GTCCACATAGAGCGAGTTTGAAGAAATCTATTAAGACGATTTCTC----- 336
QY 176 LYSSERCYSLEUYSANLSTHLYSGLUASNVALSPVALPROLEUVALILECYSGLY 195

```

```

DB 337 -----AGAGTAAAGATCGTGCATGAGTCCCATGATTTAATGCT 378
QY 196 ASNLVSGLYASPARGASPPHETRYARGGLUVALASPLNARGLUILEGLINLEUVAL 215
DB 379 AATTAAGCAGATCTGATTCATCAAAAGACAGGTAAACAGAGAAAGACAAACAGTTAGCA 438
QY 216 GLYASPSAPPROGILNARGYSAIATYRPHGLUILESERIALYLSYASNSERLEU 235
DB 439 CGGCAG-----CTTAAGGTAAACATACATGAGGACATCAGCAAAATTAGAGATGATGTA 492
QY 226 ASPLINMETPHEARGALALEUPHEALAMEALALYSLEUPROSERGIUMET----- 252
DB 493 GATCAAGCTTTCATGAACTTGCCGGTTATACGAAATTTCAAGACGAGAAATGTCCT 552
QY 253 ---SERRPROASPLEUHSISARGLYS 259
DB 553 CCTTCAACCAAGAACCAACACGAGAAA 576

RESULT 7
US-08-306-691B-15
; Sequence 15, Application US/08306691B
; Patent No. 5734039
; GENERAL INFORMATION:
; APPLICANT: Calabretta, Bruno
; TITLE OF INVENTION: OLIGONUCLEOTIDES TARGETING COOPERATING ONCOGENES
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seidel, Gonda, Lavorgna & Monaco, P. C.
; STREET: Two Penn Center, Suite 1800
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19102
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 720 KB
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/306,691B
FILING DATE: September 15, 1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Monaco, Daniel A.
REGISTRATION NUMBER: 30,480
REFERENCE/DOCKET NUMBER: 8321-8
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-8383
TELEFAX: (215) 568-5549
TELEX: No. 5734039e
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 5775 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-306-691B-15

Alignment Scores:
Pred. No.: 1,586-15 Length: 5775
Score: 279.00 Matches: 67
Percent Similarity: 54.59% Conservative: 34
Best Local Similarity: 36.22% Mismatches: 72
Query Match: 16.12% Indels: 12
DB: 1 Gaps: 4

```

TELEFAX: 516-742-4366

RESULT 11

US-08-429-964-85
 ; Sequence 85, Application US/08429964
 ; Patent No. 5962243
 ; GENERAL INFORMATION:
 ; APPLICANT: BROWN, MICHAEL S.
 ; APPLICANT: GOLDSTEIN, JOSEPH L.
 ; APPLICANT: REISS, YUVAL
 ; APPLICANT: JAMES, GUY L.
 ; TITLE OF INVENTION: METHODS FOR THE IDENTIFICATION OF FARNESYL
 ; TITLE OF INVENTION: TRANSFERASE INHIBITORS
 ; NUMBER OF SEQUENCES: 85
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: ARNOLD, WHITE & DURKEE
 ; STREET: P.O. BOX 4433
 ; CITY: HOUSTON
 ; STATE: TEXAS
 ; COUNTRY: UNITED STATES OF AMERICA
 ; ZIP: 77210

COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; OPERATING SYSTEM: IBM PC compatible
 ; SOFTWARE: Patent Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/429,964
 ; FILING DATE: 27-APR-1995
 ; CLASSIFICATION: 435

PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/021,625
 ; FILING DATE: 16-FEB-1993
 ; CLASSIFICATION: 435

APPLICATION NUMBER: US 07/822,011
 ; FILING DATE: ABANDONED
 ; CLASSIFICATION: 435

APPLICATION NUMBER: PCT/US/91/02650
 ; FILING DATE: 18-APR-1991
 ; CLASSIFICATION: 435

APPLICATION NUMBER: US 07/615,715
 ; FILING DATE: 20-NOV-1990
 ; CLASSIFICATION: 435

APPLICATION NUMBER: US 07/510,706
 ; FILING DATE: 18-APR-1990 (ABANDONED)
 ; CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
 ; NAME: PARKER, DAVID L.
 ; REGISTRATION NUMBER: 32,165
 ; REFERENCE/DOCKET NUMBER: UTSD:432/PAR

TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (512) 418-3000
 ; TELEFAX: (713) 789-2679
 ; TELEX: 79-0924

INFORMATION FOR SEQ ID NO: 85:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 607 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear

US-08-429-964-85

Alignment Scores:
 Pred. No.: 1.24e-16 Length: 607
 Score: 276.00 Matches: 66
 Percent Similarity: 54.59% Conservative: 35
 Best Local Similarity: 35.68% Mismatches: 72
 Query Match: 15.94% Indels: 12
 DB: 2 Gaps: 4

US-09-709-103-3F1 (1-332) x US-08-429-964-85 (1-607)

Qy 76 TyrArgMetValIleLeuGlySerSerLysValGlyThrAlaIleValSerArgPhe 95
 Db 10 TATAACTTGTGGTAGTTGGAGTCTGTGGCAGGAGTCTGTGGACGACGCTA 69

Qy 96 LeuThrGlyArgPheGluAspAlaTyrThrProThrIleGluAspPheHisArgLysPhe 115
 Db 70 ATTGACAATCATTTTGTGACGAATATGATCCAAATAGAGGATTCCTACAGCAAGCAA 129
 Qy 116 TyrSerIleArgGlyGluValTyrGlnLeuAspIleLeuAspThrSerGlyAsnHisPro 135
 Db 130 GTAGTAATTTGATGGAGAAACCTGCTCTTGGATATTCTCGACACAGCGGTCAAGAGGAG 189
 Qy 136 PheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIleLeuValPheSer 155
 Db 190 TACAGTGCATGAGGACCATGATCAGGAGCTGGGAGGCTTCTTGTGATTGCTCC 249
 Qy 156 LeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGlnGlnIleLeuAspThr 175
 Db 250 ATAATATATACTAAATCATTTGAAGATATTCACCATATATAGAACAAAT----- 300
 Qy 176 LysSerCysLeuLysAsnLysThrLysGluAsnValAspValProLeuValIleCysGly 195
 Db 301 -----AAAAGAGTTAAGGACTCTGAAGATGTACCTATGTCTCTAGTAGGA 345
 Qy 196 AsnLysGlyAspArgAspPheTyrArgGluValAspGlnArgGluIleGluGlnLeuVal 215
 Db 346 AATAAATGTGATTGCTCT---TCTAGAACAGTAGACACAAACAGGCTCAGGACTTAGCA 402
 Qy 216 GlyAspAspProGlnArgCysAlaTyrPheGluIleSerAlaLysLysAsnSerSerLeu 235
 Db 403 AGAAGT-----TATGGAATTCCTTTTATTCAAACATCAGCAACAGACAGAGGTGT 456
 Qy 236 AspGlnMetPheArgAlaLeuPhe---AlaMetAlaLysLeuProSerGluMetSerPro 254
 Db 457 GATGATCCCTTCTATACATTAGTTCGAGAAATTCGAAACATATAAGAAAGATGAGCAA 516
 Qy 255 AspLeuHisArgLys 259
 Db 517 GATGGTAAAAAGAAAG 531

RESULT 12

US-08-884-866A-9
 ; Sequence 9, Application US/08884866A
 ; GENERAL INFORMATION:
 ; APPLICANT: Chien, Shu
 ; APPLICANT: Shyy, John Y-J
 ; TITLE OF INVENTION: GENE THERAPY IN CORONARY ANGIOPLASTY AND
 ; TITLE OF INVENTION: BYPASS
 ; FILE REFERENCE: UCSD1100-1
 ; CURRENT APPLICATION NUMBER: US/08/884,866A
 ; PRIOR FILING DATE: 1997-06-30
 ; PRIOR APPLICATION NUMBER: 60/030,358
 ; NUMBER OF SEQ ID NOS: 25
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 9
 ; LENGTH: 480
 ; TYPE: DNA
 ; ORGANISM: Rattus norvegicus
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (49)---(51)
 ; OTHER INFORMATION: nnn = Any nucleic acid triplet, except for UCA,
 ; OTHER INFORMATION: UCC, UCG, UCU, AGC, or AGU
 US-08-884-866A-9

Alignment Scores:
 Pred. No.: 9.88e-17 Length: 480
 Score: 275.50 Matches: 63
 Percent Similarity: 56.25% Conservative: 31
 Best Local Similarity: 37.72% Mismatches: 62
 Query Match: 15.92% Indels: 11
 DB: 4 Gaps: 3

US-09-709-103-3F1 (1-332) x US-08-884-866A-9 (1-480)

```

QY 76 TyrArgMetValIleLeuGlySerSerLysValGlyThrAlaIleValSerArgPhe 95
Db 10 TATAAGCTGCTGCTGGCGCGCGCGCTGTGGCAANNNGCGCTGACCATTCAGCTG 69
QY 96 LeuThrGlyArgPheGluAspAlaTyrThrProThrIleGluAspPheIleArgLysPhe 115
Db 70 ATCCGAACCATTTTGTGTGACGATACGACCCCACTATAGAGATTCTTACCGAGACAG 129
QY 116 TyrSerIleArgGlyGluValTyrIleLeuAspIleLeuAspThrSerGlyAsnHisPro 135
Db 130 GTGGCATTTGATGGGAGACGCTGCTGTGGCATCTCGATACCGCGCGCTGAGAGAG 189
QY 136 PheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIleLeuValPheSer 155
Db 190 TACAGCGCATGCGGACCACTCAATGCGCACCGGGAGGCGCTTCTGCTGTGTGTGCC 249
QY 156 LeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGlnGlnIleLeuAspThr 175
Db 250 ATCAACACACCAAGCTTTTGTGACATCCACCACTACAGGAGACAGATC----- 300
QY 176 LysSerCysLeuLysAsnLysThrLysGluAsnValAspValProLeuValIleCysGly 195
Db 301 -----AAACGGTGAAGCACTCGGATGACCTGCCATGCTGCTGTGTGGG 345
QY 196 AsnLysGlyAspArgAspPheTyrArgGluValAspGlnArgGlnIleGluGlnLeuVal 215
Db 346 AACAAAGTGTACCTG---GCTGACAGCACTGTGATCTCGGAGGCTCAGACCTCCGC 402
QY 216 GlyAspAspProGlnArgCysAlaTyrPheGluIleSerAlaLysLysAsnSerLeu 235
Db 403 CGAAGC-----TACGGCATCCCTACATGAGACCTCGGCCAAGACCGGAGGAGTG 456
QY 236 AspGlnMetPheArgAlaLeu 242
Db 457 GAGGATGCTTCTTACAGCTTG 477

RESULT 13
US-09-167-322-12
Sequence 12, Application US/09167322
Patent No. 6365151
GENERAL INFORMATION:
APPLICANT: Allegheny University of the Health
Sciences, Halpern, Michael S.
England, James M.
TITLE OF INVENTION: CANCER VACCINE
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESS: Seidel, Conda, Lavorigna & Monaco, P.C.
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19102
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patenting Release #1.0, Version #1.30
CURRENT APPLICATION DATA: US/09167,322
APPLICATION NUMBER: 07-Oct-1998
CLASSIFICATION: <Unknown>
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: PCT/US97/00582
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Monaco, Daniel A.
REGISTRATION NUMBER: 30,480
REFERENCE/DOCKET NUMBER: 7933-33 PC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-8383
TELEFAX: (215) 568-5549
INFORMATION FOR SEQ ID NO: 12:

```

```

SEQUENCE CHARACTERISTICS:
LENGTH: 4480 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-09-167-322-12

Alignment Scores:
Pred. No.: 2,3e-15 Length: 4480
Score: 275.50 Matches: 89
Percent Similarity: 46.01% Conservative: 38
Best Local Similarity: 32.25% Mismatches: 92
Query Match: 15.92% Indels: 58
DB: Gaps: 10

US-09-709-103-3fl (1-332) x US-09-167-322-12 (1-4480)

QY 7 GlyAlaProSerProSer-----ArgAlaGlnProGlnGlnSerProPro 21
Db 1130 GACAGCTCGAGACCGAGACGCTGATCACTCGGTGAGATCAAGAGCTCTCCGAG 1189
QY 22 AlaAlaHisProAlaCysHisProSerAspProGln-ProLeuSerAlaLeuSerAl 41
Db 1190 GTC-----GGGACAGTCGCGCCGACGACGCGGTGGCGAGAGC 1228
QY 41 aProArgProProSerArgProLeuCysProMetLysLeuAlaMetIleLysLysMe 61
Db 1229 TCCT-----GGTTGGCAGCCCTCTGTGAGAGCGAT 1258
QY 61 tCysProSerAspSerGluLeuSerIleProAlaLysAsnCysTyrArgMetValIle 81
Db 1259 GACAGAA-----TACAGCTGTGTGTGT 1282
QY 81 uGlySerSerLysValGlyLysThrAlaIleValSerArgPheLeuThrGlyArgPheG 101
Db 1283 GGGCGCTAGAGGGCGTGGAAAGTGCCTGACCATCCAGCTGATCCAGAACATTTTGT 1342
QY 101 uAspAlaTyrThrProThrIleGluAspPheHisArgLysPheTyrSerIleArgGlyG 121
Db 1343 GACAGAGATGATCCCACTTATAGAGACCTCAACCGAAGACAGTATGATGATGGGA 1402
QY 121 uValTyrGlnLeuAspIleLeuAspThrSerGlyAsnHisProPheProAlaMetArg 141
Db 1403 GACGGTTACTGACATCTTATAGACAGGATCAAGAGTCAAGAAAGTATAGTCCATGGGA 1462
QY 141 gluSerIleLeuThrGlyAspValPheIleLeuValPheSerLeuAspAsnArgAsp 161
Db 1463 CCAGTACATGGCCACAGGGAGGCTTCCCTGTGATTTGGCATCAACACCAAGTC 1522
QY 161 rPheGluGluValGlnArgLeuArgGlnGlnIleLeuAspThrLysSerCysLeuLys 181
Db 1523 CTTTGAAGACATCCATCAGTACAGGAGCAAGT-----AA 1558
QY 181 nLysThrLysGluAsnValAspValProLeuValIleCysGlyAsnLysGlyAspArg 201
Db 1559 GCGGGTGAAGATTCAGATGATGTCCTGCTGTGTGGCAAGAGTGAAGCTGGC 1618
QY 201 pPheTyrArgGluValAspGlnArgGlnIleGluGlnLeuValGlyAspAspProGln 221
Db 1619 CGGTAC---ACTGTGAGTCTCGGACAGCCAGACCTTCTCGCAGC-----TATGG 1669
QY 221 gCysAlaTyrPheGluIleSerAlaLysLysAsnSerSerLeuAspGlnMetPheArg 241
Db 1670 CATCCCTCATTTGAACATCAAGCCAGACCCAGCAGTGTGAGATGCTTCAAC 1729
QY 241 alaLys-----PheAlaMetAlaLysLeu-----ProSerGluMetSe 253
Db 1730 ACATGATCGAGATTCGGACAGATTAAGTGGGAAATGAACCCGCTGATGAGATGG 1789
QY 253 rProAspLeuHisArgLysValSerValGlnTyrCysAspValLeu 268
Db 1790 CCTGG-CTGCAT-----GAGCTGCAAGTGTGTGTCTGTG 1825

```

RESULT 14

US-08-884-866A-10
 ; Sequence 10, Application US/08884866A
 ; GENERAL INFORMATION:
 ; APPLICANT: Chien, Shu
 ; APPLICANT: Shyy, John Y-J
 ; TITLE OF INVENTION: GENE THERAPY IN CORONARY ANGIOPLASTY AND
 ; FILE REFERENCE: UCSD1100-1
 ; CURRENT APPLICATION NUMBER: US/08/884,866A
 ; CURRENT FILING DATE: 1997-06-30
 ; PRIOR APPLICATION NUMBER: 60/030,358
 ; PRIOR FILING DATE: 1996-11-08
 ; NUMBER OF SEQ ID NOS: 25
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 10
 ; LENGTH: 450
 ; TYPE: DNA
 ; ORGANISM: Rattus norvegicus
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (49)...(51)
 ; OTHER INFORMATION: nnn = Any nucleic acid triplet, except for UCA,
 ; OTHER INFORMATION: UCC, UCG, UCU, AGC, or AGU
 US-08-884-866A-10

Alignment Scores:

Pred. No.:	5.94e-16	Length:	450
Score:	266.50	Matches:	61
Percent Similarity:	58.06%	Conservative:	29
Best Local Similarity:	39.35%	Mismatches:	54
Query Match:	15.40%	Indels:	11
DB:	4	Gaps:	3

US-09-709-103-3F1 (1-332) x US-08-884-866A-10 (1-450)

QY	76	TyrArgMetValIleLeuGlySerLysValGlyThrAlaIleValSerArgPhe	95
Db	10	TATAAGCTGTGTGTGGCGCGCGGTGGGCAANNNGCCTGACCATCCAGCTG	69
QY	96	LeuThrGlyArgPheGluAspAlaTyrThrProThrIleGluAspPheHisArgLysPhe	115
Db	70	ATCCAGAACCATTTGTGGACGAATACGACCCACTATAGAGATTCTTACCGAAGCAG	129
QY	116	TyrSerIleArgGlyGluValTyrGlnLeuAspIleLeuAspThrSerGlyAsnHisPro	135
Db	130	GTGGTCAATTGGGGAGACGCTCTGTGGACATCTCGGATACCGCGCGCTGGAGCAG	189
QY	136	PheProAlaMetArgLeuSerIleLeuThrGlyAspValPheIleLeuValPheSer	155
Db	190	TACAGCGCATCGGGACGACGATCAATGCGCACCGGGAGGCTTCTGTGTGTGTGGC	249
QY	156	LeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGlnGlnIleLeuAspThr	175
Db	250	ATCAACACACCAAGCTCTTTGAGGACATCCACGACATACGAGGAGCAGATC-----	300
QY	176	LysSerCysLeuLysAsnLysThrLysGluAsnValAspValProLeuValIleCysGly	195
Db	301	-----AAACGGGTGAGGACTCGGATGACGTGCGCCATGGCTGTGGTGGGG	345
QY	196	AsnLysGlyAspArgAspPheTyrArgGluValAspGlnArgGluIleGlnLeuVal	215
Db	346	AACAAGTGTGACCTG---GCTGCACGACGCTGTGGAATCTCGGAGGCTCAGGACCTCGCC	402
QY	216	GlyAspAspProGlnArgCysAlaTyrPheGluIleSerAlaLys	230
Db	403	CGAAGC-----TACGGCATCCCTATATGACATCTCGGCCAAG	441

RESULT 15

US-08-429-964-83
 ; Sequence 83, Application US/08429964
 ; Patent No. 5962243

GENERAL INFORMATION:
 ; APPLICANT: BROWN, MICHAEL S.
 ; APPLICANT: GOLDSTEIN, JOSEPH L.
 ; APPLICANT: REISS, YUVAL
 ; APPLICANT: JAMES, GUY L.
 ; TITLE OF INVENTION: METHODS FOR THE IDENTIFICATION OF FARNESYL
 ; NUMBER OF SEQUENCES: 85
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: ARNOLD, WHITE & DURKEE
 ; STREET: P.O. BOX 4433
 ; CITY: HOUSTON
 ; STATE: TEXAS
 ; COUNTRY: UNITED STATES OF AMERICA
 ; ZIP: 77210
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/429,964
 ; FILING DATE: 27-APR-1995
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/021,625
 ; FILING DATE: 16-FEB-1993
 ; CLASSIFICATION: 435
 ; APPLICATION NUMBER: US 07/822,011
 ; FILING DATE: ABANDONED
 ; CLASSIFICATION: 435
 ; APPLICATION NUMBER: PCT/US/91/02650
 ; FILING DATE: 18-APR-1991
 ; CLASSIFICATION: 435
 ; APPLICATION NUMBER: US 07/615,715
 ; FILING DATE: 20-NOV-1990
 ; CLASSIFICATION: 435
 ; APPLICATION NUMBER: US 07/510,706
 ; FILING DATE: 18-APR-1990 (ABANDONED)
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: PARKER, DAVID L.
 ; REGISTRATION NUMBER: 32,165
 ; REFERENCE/DOCKET NUMBER: UTSD:432/PAR
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (512) 418-3000
 ; TELEFAX: (713) 789-2679
 ; TELEX: 79-0924
 ; INFORMATION FOR SEQ ID NO: 83:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 574 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 US-08-429-964-83

Alignment Scores:

Pred. No.:	8.37e-16	Length:	574
Score:	266.50	Matches:	62
Percent Similarity:	53.85%	Conservative:	36
Best Local Similarity:	34.07%	Mismatches:	67
Query Match:	15.40%	Indels:	17
DB:	2	Gaps:	4

US-09-709-103-3F1 (1-332) x US-08-429-964-83 (1-574)

QY	76	TyrArgMetValIleLeuGlySerLysValGlyThrAlaIleValSerArgPhe	95
Db	10	TATAAACTTGTGTGTGGAGCTTGTGGCGTAGGCACAGCTGCTTACCATACAGCTA	69
QY	96	LeuThrGlyArgPheGluAspAlaTyrThrProThrIleGluAspPheHisArgLysPhe	115
Db	70	ATTGACAAATCATTTTGTGGACGAATATGATCCAAATAGAGGATTCTCTACAGGAGCAA	129

QY 116 TyrSer11eaGGLyGluValTyrGlnLeuAsp11eLeuAspThrSerGlyAsnHisPro 135
DB 130 GTAGTAATTGATGAGAAACCTGTCTCTTGATATTTCTCGACACAGCAGGTCAAGAGAG 189
QY 136 PheProAlaMetArgArgLeuSer11eLeuThrGlyAspValPhe11eLeuValPheSer 155
DB 190 TACAGTCGCAATAGAGGACCAACATACATGAGAGCTGGGGAGGCTTTCTTGTGATTTGCC 249
QY 156 LeuAspAsnArgAspSerPheGlnGluValGlnArgLeuArgGlnGln11eLeuAspThr 175
DB 250 ATAAATTAATTAATAATCATTTGAAAGATTTCCACCATTTATAGAGAAACAATT----- 300
QY 176 LysSerCysLeuLysAsnLysThrLysGlnAsnValAspValProLeuVal11eCysGly 195
DB 301 -----AAAGAGTTAAGGACTCTGAAGATGTACCTATGCTCTAGTGA 345
QY 196 AsnLysGlyAspArgAspPheTyrArgGluValAspGlnArgGln11eGlnLeuVal 215
DB 346 AATAAATGATTTGGCT--TCTAGACAGTAGACACAAAACAGGCTCAGGACTTAGCA 402
QY 216 GlyAspAspProGlnArgCysAlaTyrPheGlnLysSerAlaLysLysAsnSerSerLeu 235
DB 403 AGAAGT-----TATGGAATTCCTTTATTTCAAAACATCAGCAAAAGACAAAGACAGAGGTG 456
QY 236 AspGlnMetPheArgAlaLeu-----PheAlaMetAlaLysLeuPro 249
DB 457 GAGGATGCTTTTATACATTTGTGAGAGAGATCCGACAAATACGATTGAAAAAATCAGC 516
QY 250 SerGlu 251
DB 517 AAAGAA 522

Search completed: December 30, 2002, 19:44:17
Job time : 88 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 30, 2002, 18:07:41 ; Search time 90 seconds
(without alignments)

Title: US-09-709-103-3f1
Perfect score: 1731
Sequence: 1 GIPGAGAPSPRAQPEQSP.....IREKASAGSOAKDKERCVIS 332
1498.554 Million cell updates/sec

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 363474 seqs, 203117208 residues

Total number of hits satisfying chosen parameters: 726948

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlp
-O=/cgn2_1/USPFO.spool/SULLIVAN3f1/runat_30122002.144155.22191/app_query.fasta_1.519
-DB=Published Applications NA -QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LDB=0 -DOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -THR MAX=100
-MAXLEN=2000000000 -USER=SULLIVAN3f1@cgn2_1.36 @runat_30122002.144155.22191
-NCPU=6 -ICPU=3 -NO_XLPHY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				
Result No.	Score	Query Match	Length DB ID	Description
1	919.5	53.1	3082	10 US-09-778-963A-1 Sequence 1, Appli
2	540	31.2	405	10 US-09-960-352-10273 Sequence 10273, A
3	526	30.4	11221	10 US-09-778-963A-3 Sequence 3, Appli
4	376	21.7	368	10 US-09-864-761-21643 Sequence 21643, A

5	312.5	18.1	3346	9 US-09-764-868-67 Sequence 67, Appl
6	303.5	17.5	688	9 US-09-764-868-490 Sequence 490, App
7	283.5	16.4	2040	12 US-10-044-090-336 Sequence 336, App
8	278.5	16.1	197997	10 US-09-822-246-3 Sequence 3, Appli
9	276.5	16.0	570	12 US-10-104-484-1 Sequence 1, Appli
10	276.5	16.0	570	12 US-10-104-484-3 Sequence 25, Appli
11	275.5	15.9	551	10 US-09-765-298A-25 Sequence 25, Appli
12	272.5	15.7	570	10 US-09-765-298A-27 Sequence 27, Appli
13	268.5	15.5	942	10 US-09-962-832-256 Sequence 256, App
14	264.5	15.3	847	10 US-09-822-849A-40 Sequence 40, Appl
15	264.5	15.3	930	10 US-09-801-368-285 Sequence 285, App
16	262.5	15.2	1271	12 US-10-044-090-110 Sequence 110, App
17	259.5	15.0	536	10 US-09-867-701-5566 Sequence 5566, Ap
18	254	14.7	952	10 US-09-917-800A-1389 Sequence 1389, Ap
19	245.5	14.2	969	10 US-09-801-368-287 Sequence 287, App
20	241.5	14.0	1088	10 US-09-917-800A-1616 Sequence 1616, Ap
21	237	13.7	2418	9 US-09-764-868-84 Sequence 84, Appl
22	237	13.7	2479	9 US-09-764-868-69 Sequence 69, Appl
23	236.5	13.7	1085	9 US-10-067-813-3 Sequence 3, Appli
24	236	13.6	1119	9 US-10-067-813-1 Sequence 1, Appli
25	232	13.4	1191	12 US-10-044-090-111 Sequence 111, App
26	231.5	13.4	452	10 US-09-960-352-4253 Sequence 4253, Ap
27	228	13.2	1296	9 US-09-954-531-983 Sequence 983, App
28	228	13.2	1296	9 US-09-954-531-1378 Sequence 1378, Ap
29	225.5	13.0	454	10 US-09-960-352-5830 Sequence 5830, Ap
30	222.5	12.9	1224	10 US-09-972-529-1 Sequence 1, Appli
31	222.5	12.9	1152	10 US-09-788-654A-1 Sequence 1, Appli
32	222	12.8	752	9 US-10-108-605-44 Sequence 44, Appl
33	221.5	12.8	752	9 US-09-764-868-501 Sequence 501, App
34	211	12.2	1199	10 US-09-794-257-10 Sequence 10, Appl
35	210.5	12.2	1274	10 US-09-925-302-91 Sequence 91, Appl
36	208	12.0	412	10 US-09-960-352-6794 Sequence 6794, Ap
37	205	11.8	1964	10 US-09-925-297-367 Sequence 367, App
38	200.5	11.6	612	10 US-09-925-300-631 Sequence 3, Appli
39	200	11.6	1537	10 US-09-925-300-631 Sequence 4, Appli
40	198	11.4	1023	10 US-09-794-257-4 Sequence 674, App
41	194.5	11.2	836	10 US-09-925-300-678 Sequence 678, App
42	194.5	11.2	2667	10 US-09-962-832-224 Sequence 224, App
43	193	11.1	1443	10 US-09-960-352-6793 Sequence 6793, Ap
44	192	11.1	401	10 US-09-794-257-12 Sequence 12, Appl
45	192	11.1	552	

ALIGNMENTS

RESULT 1

US-09-778-963A-1
; Sequence 1, Application US/09778963A
; Patent No. US20020115172A1
; GENERAL INFORMATION:
; APPLICANT: NEELAM, Beena et al
; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
; FILE REFERENCE: CLO01112
; CURRENT APPLICATION NUMBER: US/09/778,963A
; CURRENT FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3082
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-778-963A-1

Alignment Scores:
Pred. No.: 2.01e-71
Score: 919.50
Percent Similarity: 71.39%
Best local Similarity: 56.93%
Query Match: 53.12%
DB: 10
Length: 3082
Matches: 193
Conservative: 49
Mismatch: 72
Indels: 25
Gaps: 8

US-09-709-103-3f1 (1-332) x US-09-778-963A-1 (1-3082)

```

QY      1 GlyIleProSerGlyValAlaGlyAla---ProSerProSerArgAlaGlnProGluGlnSer 19
      40 GGGGTCCTCCAGACAGCGGAGGAGATCCCGGACGATGACCCGCGAGCCACCAACACTCA 99
QY      20 ProProAlaAlaHisProAlaCysHisProSerArgProGlnProLeuSerAlaLeu--- 38
      100 CTGGAGAGCTCGCGCGCTGAGACAGCAGGACCTCCCGCAG---CTCCCGCGCTTCCA 156
QY      39 -----LeuSerAlaProArgProProSerArgProLeuCySPrometLysLeuAla 55
      157 GGCAGCTCTTGTAGCGCTGACAGGAGCCGCGCGCATTTCCACAGCC-----CGA 207
QY      56 AlaMetIleLysLysMetCysProSerArgSerGluLeuSerIleProAlaLysAsnGly 75
      208 GCCATGATGAAGACTTGTTCACCGGAACTGCACGCTCAGTGTGCGCGCAAAACTCA 267
QY      76 TyrArgMetValIleLeuGlySerSerLysValGlyLysThrAlaIleValSerArgPhe 95
      268 TACCGCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 327
QY      96 LeuThrGlyArgPheGluAspAlaIleTyrThrProThrIleGluAspPheHisArgLysPhe 115
      328 CTCAATGCGCGCTTGTAGACACAGTACACACCCATCGAGGACTTCCACCGGTAGGTA 387
QY      116 TyrSerIleArgGlyIleValIleTyrGlnLeuAspIleLeuAspThrSerGlyAsnHisPro 135
      388 TACAACATCCCGCGCGCATGTATACACCTGATCTGTATACCTGTGGACCAACACCCC 447
QY      136 PheProAlaMetArgArgLysLeuSerIleLeuThrGlyAspValPheIleLeuValPheSer 155
      448 TTCCCGCCATGCGGAGCTGTCTCTCACTCAAGGGAGTGTCTTCACTCGGTGTCAAC 507
QY      156 LeuAspAsnArgSerPheGluGluValGlnArgLeuArgGlnGlnIleLeuAspThr 175
      508 CTGAGTACCGGAGTCTTGTATGAGGTCAAGCGCTTCAAGAGCATCTCGAGAGTTC 567
QY      176 LysSerCysLeuLysAsnLysThrLysGluAsnValAspValProLeuValIleCysGly 195
      568 AAGTCTCTCTGAAACAAACAGACAGAGAGGCGGCGGAGCTGCCCATGTGTGTGTGT 627
QY      196 AsnLysGlyAspArg---AspPheTyrArgGluValAspGlnArgGluIleGluGlnLeu 214
      628 AACAAAGACGACGAGCGAGCTGTGCGCGAGGCTCCACACCGAGCGGAGCTGTG 687
QY      215 ValGlyAspAspProGlnArgCysAlaIleTyrPheGluIleSerAlaLysLysAsnSerSer 234
      688 GTGTCTCGGCGAC---GAGAACTCCGCTTCTGTGAGGTGTGCGCGCAGAGAAACACCAAC 744
QY      235 LeuAspGluMetPheArgAlaLeuPheAlaMetAlaLysLeuProSerGluMetSerPro 254
      745 GTGACAGAGATGTTCTACGCTGCTTCAAGCATGGCCCAAGCTGCCACAGATGAGCCCC 804
QY      255 AspLeuHisArgLysValSerValGlnTyrCysAspValLeuHisLysLys-----Ala 272
      805 GCCCTCATGCGAAGATCTCCGTGCGATGAGTGAAGCTTCCACCCAGCGCCCTTCTGC 864
QY      273 LeuArgAsnLysLysLeuLeuArgAlaGlySerGlyGlyGlyGlyAspProGlyAsp 292
      865 ATGCGCGCGCTCAAGAGATG-----GAC 888
QY      293 AlaPheGlyIleValAlaProPheAlaArgArgProSerValHisSerAspLeuMetYr 312
      889 GCCATGAGCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 948
QY      313 IleArgGlyLysAlaSerAlaGlySerGlnAlaLysAspLysGluArgCysValIle 331
      949 ATCAAGGCGCAAGGTCTTGTGGAAAGCGCAAGCGCGTGAAGGAGCAAGTGCACATC 1005

```

```

; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO: 10273
; LENGTH: 405
; TYPE: DNA
; ORGANISM: Bos taurus
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (3), (23), (31)
; OTHER INFORMATION: unsure at all n locations
; OTHER INFORMATION: Clone ID: 44-LIB3058-050-Q1-K1-C8
US-09-960-352-10273

```

```

Alignment Scores:
Pred. No.: 2,286-39 Length: 405
Score: 540.00 Matches: 108
Percent Similarity: 85.61% Conservative: 5
Best Local Similarity: 81.82% Mismatches: 18
Query Match: 31.20% Indels: 1
DB: Gaps: 0

```

US-09-709-103-3f1 (1-332) x US-09-960-352-10273 (1-405)

```

QY      32 ProGlnProLeuSerAlaLeuLeuSerAlaProArgPro-ProSerArgProLeuCySPR 51
      11 CCCCAACCGGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 70
QY      51 OmetLysLeuAlaAlaMetIleLysLysMetCysProSerArgSerGluLeuSerIlePr 71
      71 AATGAACGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 130
QY      71 AlaLysAsnCysTyrArgMetValIleLeuGlySerSerLysValGlyLysThrAlaI 91
      131 GGCAGAAAGACGTCAACGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 190
QY      91 eValSerArgPheLeuThrArgLysArgPheGluAspAlaIleTyrThrProThrIleGluAspPh 111
      191 GGTGTGCGCTTCTGTGAGCGGCGCTACAGAGACGCTTACACGCCACCATCGAAGACTT 250
QY      111 eHisArgLysPheTyrSerIleArgGlyIleValIleTyrGlnLeuAspIleLeuAspThrSe 131
      251 CCACCGCAAGTTCTACATGCTATTCAGAGCGAGAGGTACACAGCTGACATCTCTGACACGTA 310
QY      131 rGlyAsnHisProPheProAlaMetArgArgLysSerIleLeuThrGlyAspValPheI 151
      311 CGGCAACCGCGCATTTCCCGCGCATGCGACGCTTGTGATCTTACCGGAGACGTGTTCAT 370
QY      151 eLeuValPheSerLeuAspAsnArgAspSerPhe 162
      371 CTTAGGGTTTCAAGACGTACAACCGGCACTCTTC 404

```

```

RESULT 3
US-09-778-963A-3
; Sequence 3, Application US/09778963A
; Patent No. US20020115172A1
; GENERAL INFORMATION:
; APPLICANT: NEELAM, Beena et al
; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS, THESE HUMAN RAS-LIKE
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING
; FILE REFERENCE: CL001112
; CURRENT APPLICATION NUMBER: US/09/778, 963A
; CURRENT FILING DATE: 2001-02-08

```

```
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 11221
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-778-963A-3

Alignment Scores:
Pred. No.: 2,986-36 Length: 11221
Score: 526.00 Matches: 106
Percent Similarity: 72.40% Conservative: 37
Best Local Similarity: 55.21% Mismatches: 33
Query Match: 30.39% Indels: 16
DB: 10 Gaps: 4

US-09-709-103-3F1 (1-332) x US-09-778-963A-3 (1-11221)

QY 143 SerIleLeuThrGlyAspValPheIleLeuValPheSerLeuAspAsnArgSerPhe 162
Db 7681 TCTCTCCCTGCGAGGGATGCTTCTCATCTCTGTTTTCAGCTGGATAACCGGAGTCTCTC 7740

QY 163 GluGluValGlnArgLeuArgGlnGlnIleLeuAspThrIlySerCysLeuIlyAsnLys 182
Db 7741 GATGAGGTCAAGCGCTTCAGAGCAGATCTCGAGGTCAAGTCTCGCTGAGAAACAAG 7800

QY 183 ThrLysGluAsnValAspValProLeuValIleCysGlyAsnLysGlyAspArg---Asp 201
Db 7801 ACCAAGGAGCGCGCGAGTGCCTCATCTGTGCAACAAGACGACCGCGGAG 7860

QY 202 PheTyrArgGluValAspGlnArgGluIleGluGlnLeuValGlyAspAspProGlnArg 221
Db 7861 CTGTCCCGCAGGTCCCAACCGAGCGCGAGTGTGTGTGTGCGCGGAC--GAGAAC 7917

QY 222 CysAlaTyrPheGluIleSerAlaLysLysAsnSerSerLeuAspGlnMetPheArgAla 241
Db 7918 TCGCGCTACTTCGAGGTGTCGCCAAGAGAACACCAACAGTGGACGAGATGTTCTACGTG 7977

QY 242 LeuPheAlaMetAlaLysLeuProSerGluMetSerProAspLeuHisArgLysValSer 261
Db 7978 CTCCTTCAGCATGGCCAGCTGCCACAGATGAGCGCCCGCTGCATCGCAAGATCTCC 8037

QY 262 ValGlnTyrCysAspValLeuHisLysLys-----AlaLeuArgAsnLysLysLeuLeu 279
Db 8038 GTGCAGTACGTGAGCGCTTCCACCGAGCGCTTCTGCAATGCGCGCGCGTCAAGGAGATG 8097

QY 280 ArgAlaGlySerGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 299
Db 8098 -----GAGCGCTATGGCATGTGCTCGCC 8121

QY 300 PheAlaArgArgProSerValHisSerAspLeuMetTyrIleArgGluLysAlaSerAla 319
Db 8122 TTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 8181

QY 320 GlySerGlnAlaLysAspLysGluArgCysValIle 331
Db 8182 GAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 8217

RESULT 4
US-09-864-761-21643
; Sequence 21643, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeonica-X-1
; CURRENT APPLICATION NUMBER: US/09/864, 761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180, 312
```

```
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO 21643
; LENGTH: 368
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL022334.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.8
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.2
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.9
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.9
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3
; OTHER INFORMATION: NT HIT: g11418034, EVALUATE 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: P13856, EVALUATE 8.00e-12
; OTHER INFORMATION: EST_HUMAN HIT: BE389944.1, EVALUATE 0.00e+00
US-09-864-761-21643

Alignment Scores:
Pred. No.: 4,51e-25 Length: 368
Score: 376.00 Matches: 79
Percent Similarity: 67.91% Conservative: 12
Best Local Similarity: 58.96% Mismatches: 28
Query Match: 21.72% Indels: 15
DB: 10 Gaps: 2

US-09-709-103-3F1 (1-332) x US-09-864-761-21643 (1-368)

QY 16 ProGluGlnSerProAlaAlaHisProAlaCysHisPro-----SerAspPro-G1 33
Db 4 CCTAGACAGAGCGCTAGAGAGCGCCAGAGGTGGTGGGCGCCAGGTCTCGGGGGGCGCCCTGAT 63

QY 33 nProLeuSerAlaLeuLeuSerAlaProArgProArgProSerArgProLeuCysProMetLys 53
```


QY 19 SerProProAlaHisProAlaCysHisProSerAspProGlnProLeuSerAlaLeu 38
 Db 14 TCCCTCCCTCCCT-----CCTGTAACCTCTCTCTTCCCGGTTGGCCCTCTGCTCCT 67
 QY 39 LeuSerAlaPro-----ArgProProSerArgProLeuCysProMet 52
 Db 68 CTGACCTCTCTCCCGAGTTCGAGATTCGCGCCACCTTCGCTCTAGCGCGGCC 127
 QY 53 LysLeuAlaAlaMetLeuLysLysMetCysProSerAspSerGluLeu---SerIlePro 71
 Db 128 ACAGCTAGCGGGGTGATC-----ITTCCTCCCTCTCGTAGGAGTTGGTGAAGTGAGA 181
 QY 72 AlaLysAsnCysTyArgMetValLleLeuLysSerLysValGlyLysThraLle 91
 Db 182 CTCATGAGGAATACAGAGTAGTGTGTAGGGAGTGGAGGGTGGCAATCTGCCCTT 241
 QY 92 ValSerArgPheLeuThrGlyArgPheGluAspAlaTyThrProThrIleGluAspPhe 111
 Db 242 ACTGTGCAGTTGTCACTGGGACTTTTCTGAGAAATATGACCCACCATTTGAAGATTTC 301
 QY 112 HisArgLysPheTySerIleArgGlyGluValTyGlnLeuAspIleLeuAspThrSer 131
 Db 302 TACCCCAAGAGATCGAAGTGGACTCTTCCCTCCGCTCGTGGAAATCTGGACCGCA 361
 QY 132 GlyAsnHisProPheProAlaMetArgLeuSerIleLeuThrGlyAspValPheIle 151
 Db 362 GGAAGTGCAGTTGTGCTCCATGAGATCTTACATCAAAACGGCCAAAGTTTCATC 421
 QY 152 LeuValPheSerLeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGlnGln 171
 Db 422 CTGTTTATAGCTTGTATCAACAGCTTTTTCAGATATCAAGCCCATGAGAGATCAA 481
 QY 172 IleLeuAspThrLysSerCysLeuLysAsnLysThrLysGluValAsnValProLeu 191
 Db 482 ATTGTC-----AGAGTGAAGAGATATGAAAAAGTCCCACTA 517
 QY 192 ValLleCysGlyAsnLysGlyAspArgAspPheTyArgGluValAspGlnArgGluIle 211
 Db 518 ATCCAGTAGGAATAAAGTGTGATCTGGAACCAAGAGAGGTATGCTTTCAGAAGGC 577
 QY 212 GluGlnLeuValGlyAspAspProGlnArgCysAlaTyThrPheGluIleSerAlaLysLys 231
 Db 578 ANAGCTCTGGCTCAAGAA-----TGGGGTGTCTTATGAGACATCGGCNAAAAGT 631
 QY 232 AsnSerSerLeuAspGlnMetPhe 239
 Db 632 AAATCAATGGTGGATGAATTTT 655

RESULT 7

US-10-044-090-336
 ; Sequence 336, Application US/10044090
 ; Patent No. US20020137081A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Olga Sandman
 ; TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
 ; FILE REFERENCE: PA-0028 US
 ; CURRENT APPLICATION NUMBER: US/10/044,090
 ; CURRENT FILING DATE: 2002-01-09
 ; NUMBER OF SEQ ID NOS: 850
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 336
 ; LENGTH: 2040
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; OTHER INFORMATION: Incyte ID No. US20020137081A1 3110662CB1
 US-10-044-090-336

Alignment Scores:

Pred. No.: 5,29e-16 Length: 2040
 Score: 283.50 Matches: 66
 Percent Similarity: 55.21% Conservative: 40

Best Local Similarity: 34.38% Mismatches: 69
 Query Match: 16.38% Indels: 17
 DB: 12 Gaps: 3

US-09-709-103-3f1 (1-332) x US-10-044-090-336 (1-2040)

QY 76 TyrArgMetValLleLeuGlySerSerLysValGlyLysThrAlaIleValSerArgPhe 95
 Db 202 TATAAGCTAGTCTTCTTGGCTCAGGAGCGCTTGAAGCTGCTTGTGACTGTACAAATT 261
 QY 96 LeuThrGlyArgPheGluAspAlaTyThrProThrIleGluAspPheHisArgLysPhe 115
 Db 262 GTTCAGGAATTTTGTAGAAAAATACGATCTTACAGTAGAAGATTCTTATAGAAAGCAA 321
 QY 116 TyrSerIleArgGlyGluValTyGlnLeuAspIleLeuAspThrSerGlyAsnHisPro 135
 Db 322 GTTGAAGTAGATGCACAACAGTGTATCTTGAATCTTGGATCTGCAGGAACGGAGCAA 381
 QY 136 PheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIleLeuValPheSer 155
 Db 382 TTTACAGCAATGAGGATTTATACATGAAAAATGGACAGGATTTCGATTAGTTTATTC 441
 QY 156 LeuAsnArgAspSerPheGluGluValGlnArgLeuArgGlnGlnIleLeuAspThr 175
 Db 442 ATCAGCAGCAGTCCACATTTAAGATTACAGACTTACAGACTGAGAGACAGATTCTT 495
 QY 176 LysSerCysLeuLysAsnLysThrLysGluAsnValAspValProLeuValLleCysGly 195
 Db 496 -----CGAGTTAAAGACACATGATGATGTTCCATGATCTTGTGTGT 537
 QY 196 AsnLysGlyAspArgAspPheTyArgGluValAspGlnArgGluIleGluLeuVal 215
 Db 538 AATAAGTGTGACTTGGAA-----GATGAAAGAGTTGTAGGAAGGAACAA 582
 QY 216 GlyAspAspProGlnArg-----CysAlaTyThrPheGluIleSerAlaLysLys 231
 Db 583 GGTCAAAATCTAGCAAGACAACTGGAACACTGTGCATCTTGAATCTTCTGCAAAATCA 642
 QY 232 AsnSerSerLeuAspGlnMetPheArgAlaLeuPheAlaMetAlaLysLeuProSerGlu 251
 Db 643 AAATATAATTTAATGAGATCTTTATGACCTGTAGTCGCGCAAAATTAACAGAAAAACTCCA 702
 QY 252 MetSerProAspLeuHisArgLysValSerValGln 263
 Db 703 GTGCCTGGGAAGCTCGCAAAAGATCATCTCTCAG 738

RESULT 8

US-09-822-246-3/c
 ; Sequence 3, Application US/09822246
 ; Patent No. US20020142383A1
 ; GENERAL INFORMATION:
 ; APPLICANT: MERKULOV, Gennady et al.
 ; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
 ; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
 ; FILE REFERENCE: AND USES THEREOF
 ; FILE REFERENCE: CL001149
 ; CURRENT APPLICATION NUMBER: US/09/822,246
 ; CURRENT FILING DATE: 2001-04-02
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 3
 ; LENGTH: 197997
 ; TYPE: DNA
 ; ORGANISM: Human
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(197997)
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-822-246-3

Alignment Scores:

Pred. No.: 5.84e-13 Length: 197997
 Score: 278.50 Matches: 65

Percent Similarity: 55.21% Conservative: 41
 Best Local Similarity: 33.85% Mismatches: 69
 Query Match: 16.09% Indels: 17
 DB: 10 Gaps: 3

US-09-709-103-3fl (1-332) x US-09-822-246-3 (1-197997)

QY 76 TyrArgMetValIleLeuGlySerSerIyValGlyLysThrAlaIleValSerArgPhe 95
 Db 43187 TATAAGCTAGTCGCTTGTGCTCACAGAGCGCTTGGAAGCTGCTTGTGCTTACAAATT 43128
 QY 96 LeuThrGlyArgPheGluAspAlaTyrThrProThrIleGluAspPheHisArgLysPhe 115
 Db 43127 GTTCAGAGCAATTTTGTAGAAAATACGATCTTACGATAGAGATCTTATAGAGACAA 43068
 QY 116 TySerIleArgGlyGluValTyrGlnLeuAspIleLeuAspThrSerGlyAsnHisPro 135
 Db 43067 GTTGAAGTAGATGCACACAGCTGTATCTTGAATCTTGGATCTACAGAACGAGACAA 43008
 QY 136 PheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIleLeuValPheSer 155
 Db 43007 TTTACAGCAATAGAGGATTTTATACATGAAAAATGACACAGGTTTGCTTATGTTTTC 42948
 QY 156 LeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGlnIleLeuAspThr 175
 Db 42947 ATCAGACAGACAGTCCACTTTTACGATTTACAGACCTGAGAGAACAGATTCTT 42894
 QY 176 LysSerCysLeuYsAsnLysThrIyGluAsnValAspValProLeuValIleCysGly 195
 Db 42893 -----CGAGTTAAAGACACCTGATATGTTCCAAAGATTCTGTGGT 42852
 QY 196 AsnLysGlyAspArgAspPheTyrArgGluValAspGlnArgGlnIleGlnLeuVal 215
 Db 42851 AATAAGCTGACTTGCA-----GATGAAGAGTTGTAAGGGAAGAACAA 42807
 QY 216 GlyAspAspProGlnArg-----CysAlaTyrPheGluIleSerAlaLysLys 231
 Db 42806 GGTCAAAATCTAGCAAGCAATGAAACAACGTGCTCTTGAATCTTCTGCAAAATCA 42747
 QY 232 AsnSerLeuAspGlnMetPheArgAlaLeuPheAlaMetAlaLysLeuProSerGlu 251
 Db 42746 AAAATAAATGTTAATGAGTCTTTATATGACCTAGTGGCGCAATTACAGAAAACTCCA 42687
 QY 252 MetSerProAspLeuHisArgLysValSerValGln 263
 Db 42686 GTGCTTGGAGAGCTCGCAAAAGTCATCATGTCTAG 42651

RESULT 9
 US-10-104-484-1
 ; Sequence 1, Application US/10104484
 ; Patent No. US20020150566A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Guan, Kun-Lian
 ; APPLICANT: Stewart, Scott
 ; TITLE OF INVENTION: METHOD OF INHIBITING CANCEROUS CELL PROLIFERATION USING RAS MUTAN
 ; FILE REFERENCE: UMICH-0010
 ; CURRENT APPLICATION NUMBER: US/10/104,484
 ; CURRENT FILING DATE: 2002-03-25
 ; PRIOR APPLICATION NUMBER: 60/277,959
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 1
 ; LENGTH: 570
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Nucleotide sequence of the GDP-bound Ras mutant, RasN17N69.
 US-10-104-484-1
 Alignment Scores: 4.07e-16 Length: 570
 Pred. No.:

Score: 276.50 Matches: 72
 Percent Similarity: 52.74% Conservative: 34
 Best Local Similarity: 35.82% Mismatches: 73
 Query Match: 15.97% Indels: 23
 DB: 12 Gaps: 6

US-09-709-103-3fl (1-332) x US-10-104-484-1 (1-570)

QY 76 TyrArgMetValIleLeuGlySerSerIyValGlyLysThrAlaIleValSerArgPhe 95
 Db 10 TATAAGCTAGTCGCTTGTGCTCACAGAGCGCTTGGAAGCTGCTTGTGCTTACAAATT 69
 QY 96 LeuThrGlyArgPheGluAspAlaTyrThrProThrIleGluAspPheHisArgLysPhe 115
 Db 70 ATCCAGAACCATTTTGTGAGCAATACGACCCACTATAGAGATCTTACCGAAGCAG 129
 QY 116 TySerIleArgGlyGluValTyrGlnLeuAspIleLeuAspThrSerGlyAsnHisPro 135
 Db 130 GTGCTCATTTAGTGGAGACAGCTGCTGTTGACATCTTGATACCGCGCCAGAGAGAG 189
 QY 136 PheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIleLeuValPheSer 155
 Db 190 TACAGCCCATGCGGGAACAGTACATGCGCCAGGAGGCTTCTGTGTGTTGCC 249
 QY 156 LeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGlnIleLeuAspThr 175
 Db 250 ATCAACAACACCAAGTCTTTTGTGACATCCACAGTACAGGAGACAGATC----- 300
 QY 176 LysSerCysLeuYsAsnLysThrIyGluAsnValAspValProLeuValIleCysGly 195
 Db 301 -----AAACGGTGAAGAGCTCGAGTACGTCCTCCATGTGTGTGGGG 345
 QY 196 AsnLysGlyAspArgAspPheTyrArgGluValAspGlnArgGlnIleGlnLeuVal 215
 Db 346 AACAGGTGACCTG--GCTGACAGCACTGTGAATCTCGCAGGCTCAGACCTCGCC 402
 QY 216 GlyAspAspProGlnArgCysAlaTyrPheGluIleSerAlaLysLysAsnSerLeu 235
 Db 403 CGAAGC-----TACGGCATCCCTTACATCAGACCTCGGCCAAGACCGCGAGGAGTG 456
 QY 236 AspGlnMetPheArgAlaLeu-----PheAlaMetAlaLysLeu 248
 Db 457 GAGGATGCTCTTACAGCTTGCTGCTGAGATCCGCGACAGCAAGCTGGGAGCTGAAC 516
 QY 249 ---ProSerGlnMetSerProAspLeuHisArgLysValSerValGlnTyrCysAspVal 267
 Db 517 CCTCTGATGAGAGTGGCCCGG--CTGCAT-----GAGCTGCAAGTGTGTCTCTC 566

RESULT 10
 US-10-104-484-3
 ; Sequence 3, Application US/10104484
 ; Patent No. US20020150566A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Guan, Kun-Lian
 ; APPLICANT: Stewart, Scott
 ; TITLE OF INVENTION: METHOD OF INHIBITING CANCEROUS CELL PROLIFERATION USING RAS MUTAN
 ; FILE REFERENCE: UMICH-0010
 ; CURRENT APPLICATION NUMBER: US/10/104,484
 ; CURRENT FILING DATE: 2002-03-25
 ; PRIOR APPLICATION NUMBER: 60/277,959
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 3
 ; LENGTH: 570
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:

OTHER INFORMATION: Nucleotide sequence of the wild type Ras mutant, RasN17N69.
US-10-104-484-3

Alignment Scores:

Pred. No.: 4, 07e-16 Length: 570
Score: 276.50 Matches: 72
Percent Similarity: 52.74% Conservative: 34
Best Local Similarity: 35.82% Mismatches: 73
Query Match: 15.97% Indels: 23
DB: 12 Gaps: 6

US-09-709-103-3F1 (1-332) x US-10-104-484-3 (1-570)

```

QY 76 TyrArgMetValIleLeuGlySerSerLysValGlyLysThrAlaIleValSerArgPhe 95
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 10 TATACCTGGTGGTGGCGCGCGCGGTGGGCAAGATGCGCTGACCATCCAGCTG 69
   ::
QY 96 LeuThrGlyArgPheGluAspAlaTyrThrProThrIleGluAspPheHisArgLysPhe 115
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 70 ATCCAGAACCATTTTGTGGACCAATACGACCCCACTATAGAGGATTCCTACCGGAAGCAG 129
   ::
QY 116 TyrSerIleArgGlyGluValTyrGlnLeuAspIleLeuAspThrSerGlyAsnHisPro 135
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 130 GTGGTCATTGATGGGAGACGTGCTGTGGACATCCTGGATACCGCGCGCAGGAGGAG 189
   ::
QY 136 PheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIleLeuValPheSer 155
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 190 TACAGCGCATCGGAAACCACTATACGTACCGCGCGGAGGCGCTTCTGTGTGTGTTGCC 249
   ::
QY 156 LeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGlnGlnIleLeuAspThr 175
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 250 ATCAACAAACCAAGTCTTTTGGAGACATCCACAGTACAGGAGCAGATC----- 300
   ::
QY 176 LysSerCysLeuLysAsnLysThrLysGluAsnValAspValProLeuValIleCysGly 195
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 301 -----AAACGGGTGAAGACTCGGATGACGTGCCATGCTGTGGTGGGG 345
   ::
QY 196 AsnLysGlyAspArgAspPheTyrArgGluValAspGlnArgGluIleGlnLeuVal 215
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 346 AACAAAGTGTGACCTG---GCTCGACGGCACTGTGGAACTCTCGGACGCTCAGGACCTCGCC 402
   ::
QY 216 GlyAspAspProGlnArgCysAlaTyrPheGluIleSerAlaLysLysAsnSerSerLeu 235
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 403 CGAAGC-----TACGGCATCCCTACATCGAGACCTCGGCCCAAGACCGGAGGAGTG 456
   ::
QY 236 AspGlnMetPheArgAlaLeu-----PheAlaMetAlaLysLeu--- 248
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 457 GAGGATGCCCTTCTACAGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 516
   ::
QY 249 ---ProSerGluMetSerProAspLeuHisArgLysValSerValGlnTyrCysAspVal 267
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 517 CCTCCTGATGAGAGTGGCGCCCGG-CTGCAT-----GAGCTCAAGTGTGTGCTCTC 566
   |||
QY 268 Leu 268
   |||
Db 567 CTG 569

```

RESULT 11

US-09-765-298A-25

Sequence 25, Application US/09765298A

Patent No. US20020137017A1

GENERAL INFORMATION:

APPLICANT: ARONHEIM, AMI

TITLE OF INVENTION: METHOD FOR DETECTION PROTEIN-PROTEIN INTERACTIONS AND A KIT THERE

FILE REFERENCE: 108387.01

CURRENT APPLICATION NUMBER: US/09/765,298A

PRIOR FILING DATE: 2001-01-22

PRIOR FILING DATE: 1998-07-22

PRIOR FILING DATE: 1999-01-12

NUMBER OF SEQ ID NOS: 31

SOFTWARE: PatentIn version 3.1

SEQ ID NO 25

LENGTH: 551

TYPE: DNA

ORGANISM: Homo sapiens

US-09-765-298A-25

Alignment Scores:

Pred. No.: 4, 77e-16 Length: 551
Score: 275.50 Matches: 67
Percent Similarity: 53.48% Conservative: 33
Best Local Similarity: 35.83% Mismatches: 68
Query Match: 15.92% Indels: 19
DB: 10 Gaps: 5

US-09-709-103-3F1 (1-332) x US-09-765-298A-25 (1-551)

```

QY 76 TyrArgMetValIleLeuGlySerSerLysValGlyLysThrAlaIleValSerArgPhe 95
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 4 TATAAGCTGGTGGTGGCGCGCGGTGGGCAAGATGCGCTGACCATCCAGCTG 63
   ::
QY 96 LeuThrGlyArgPheGluAspAlaTyrThrProThrIleGluAspPheHisArgLysPhe 115
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 64 ATCCAGAACCATTTTGTGGACCAATACGACCCCACTATAGAGGATTCCTACCGGAAGCAG 123
   ::
QY 116 TyrSerIleArgGlyGluValTyrGlnLeuAspIleLeuAspThrSerGlyAsnHisPro 135
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 124 GTGGTCATTGATGGGAGACGTGCTGTGGACATCCTGGATACCGCGCGCAGGAGGAG 183
   ::
QY 136 PheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIleLeuValPheSer 155
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 184 TACAGCGCATCGGAAACCACTATACGTACCGCGCGGAGGCGCTTCTGTGTGTGTTGCC 243
   ::
QY 156 LeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGlnGlnIleLeuAspThr 175
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 244 ATCAACAAACCAAGTCTTTTGGAGACATCCACAGTACAGGAGCAGATC----- 294
   ::
QY 176 LysSerCysLeuLysAsnLysThrLysGluAsnValAspValProLeuValIleCysGly 195
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 295 -----AAACGGGTGAAGACTCGGATGACGTGCCATGCTGTGTGTTGGG 339
   ::
QY 196 AsnLysGlyAspArgAspPheTyrArgGluValAspGlnArgGluIleGlnLeuVal 215
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 340 AACAAAGTGTGACCTG---GCTCGACGGCACTGTGGAACTCTCGGACGCTCAGGACCTCGCC 396
   ::
QY 216 GlyAspAspProGlnArgCysAlaTyrPheGluIleSerAlaLysLysAsnSerSerLeu 235
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 397 CGAAGC-----TACGGCATCCCTACATCGAGACCTCGGCCCAAGACCGGAGGAGTG 450
   ::
QY 236 AspGlnMetPheArgAlaLeu-----PheAlaMetAlaLysLeu--- 248
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 451 GAGGATGCCCTTCTACAGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 510
   |||
QY 249 ---ProSerGluMetSerPro 254
   |||
Db 511 CCTCCTGATGAGAGTGGCGCCC 531

```

RESULT 12

US-09-765-298A-27

Sequence 27, Application US/09765298A

Patent No. US20020137017A1

GENERAL INFORMATION:

APPLICANT: ARONHEIM, AMI

TITLE OF INVENTION: METHOD FOR DETECTION PROTEIN-PROTEIN INTERACTIONS AND A KIT THERE

FILE REFERENCE: 108387.01

CURRENT APPLICATION NUMBER: US/09/765,298A

PRIOR FILING DATE: 2001-01-22

PRIOR FILING DATE: 1998-07-22

PRIOR FILING DATE: 1999-01-12

NUMBER OF SEQ ID NOS: 31

SOFTWARE: PatentIn version 3.1

SEQ ID NO 27

Sequence 40, Application US/09822849A
 Patent No. US20020045170A1
 GENERAL INFORMATION:
 APPLICANT: Wong, Gordon G.
 APPLICANT: Clark, Hilary
 APPLICANT: Fichtel, Kim
 APPLICANT: Agostino, Michael J.
 APPLICANT: Howes, Steven H.
 APPLICANT: Resnick, Richard J.
 APPLICANT: Gulukota, Kamalakara
 APPLICANT: Graham, James R.
 APPLICANT: Genetics Institute, Inc.
 TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
 FILE REFERENCE: GIN 6403
 CURRENT APPLICATION NUMBER: US/09/822,849A
 PRIOR FILING DATE: 2001-09-04
 PRIOR APPLICATION NUMBER: 60/195,582
 NUMBER OF SEQ ID NOS: 598
 SOFTWARE: Patent In Ver. 2.0
 SEQ ID NO 40
 LENGTH: 847
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-09-822-849A-40

Alignment Scores:
 Pred. No.: 7,69e-15 Length: 847
 Score: 264.50 Matches: 82
 Percent Similarity: 45.05% Conservative: 50
 Best Local Similarity: 27.99% Mismatches: 110
 Query Match: 15.28% Indels: 52
 DB: 10 Gaps: 7

US-09-709-103-3f1 (1-332) x US-09-822-849A-40 (1-847)

Qy 4 SerGlyAlaGlyAlaProSerProSerArgAlaGlnProGluInSerProProAlaAla 23
 Db 790 TCCCTCCCTCCCTCCAGCCGCCAGGCTCCCGCCGCCAGCCGTCAGACTCCTCTCCGACC 731
 Qy 24 HisProAlaCysHisProSerApproGlnProLeuSerAlaLeuSerAlaProArg 43
 Db 730 GCTCCCGCGCGGGGCTCCAGCGCAGCAGGACCGAGTACCTCCGCGCGGAGCCAGC 671
 Qy 44 ProSerArgPro-----LeuCyProMet----- 52
 Db 670 AG-CCGCGGCTCCGAGGCGCTCCGCGCGCGGAGCTGGCTCCGCTGCAGATTCTTCTTA 612
 Qy 53 -----LysLeuAlaAlaMetIleLysLysMetCysProSer 64
 Db 611 ATCCTTTGGTGAACACTGAGACACAAATGGCTGCATAATAGCCCAAG-----GGT 561
 Qy 65 AspSerGluLeuSerIleProAlaLysAsnCysTyrArgMetValIleLeuGlySerSer 84
 Db 560 CAGAAATCTTGGCTTTA-----CACAAAGTCATCATGGTGGCGAGTGGT 516
 Qy 85 LysValGlyLysThralalleValSerArgPheLeuThrGlyArgPheGluAspAlaTyr 104
 Db 515 GCGCTGGGCAAGTCAGCTCTGACTCTACAGTTCATGTACGATGAGTTGTGGAGGACTAT 456
 Qy 105 ThrProThrIleGluAspPheHisArgLysPheTyrSerIleArgGlyGluValTyrGln 124
 Db 455 GAGCTTACCAAGCAGACAGCTATCCGGAAGAGGTAGTGTCTAGATGGGAGGAGTCCAG 396
 Qy 125 LeuAspIleLeuAspThrSerGlyAsnHisProPheProAlaMetArgArgLeuSerIle 144
 Db 395 ATCGATATCTTAGATACAGCTGGGAGGAGGAGTACCGTGCATATGAGACAACTACTTC 336
 Qy 145 LeuThrGlyAspValPheIleLeuValPheSerLeuAspAsnArgAspSerPheGluGlu 164
 Db 335 CGAAGTGGGAGGGTCTCTCTGTGTTTCTTCTATTACAGAAATGGAAATCCTTTGAGCT 276
 Qy 165 ValGlnArgLeuArgGlnGlnIleLeuAspThrLysSerCysLeuLysAsnLysThrLys 184

Db 275 ACAGCTGACTTCAGGAGCAGATTTA-----AGAGTAAAA 240
 Qy 185 GluAnValAspValProLeuValIleCysGlyAsnLysGlyAspArgAspPheTyrArg 204
 Db 239 GAAGATGAGAATGTTTCCATTCTACTGGTTGTAACAAATCAGATTTTAGAGATAAAGA 180
 Qy 205 GluValAspGlnArgGluIleGluGlnLeuValGlyAspProGlnArgCysAlaTyr 224
 Db 179 CAGGTTTCTGTAGAGAGGCAAAACAGAGCT-----GAGCAGTGCATTTGTTAACTAC 126
 Qy 225 PheGluIleSerAlaLysLysAsnSerSerLeuAspGlnMetPheArgAlaLeuPheAla 244
 Db 125 GTGGAACATCTGCTAAACACAGAGCTAATGTTGACAAAGTATTTTT-----78
 Qy 245 MetAlaLysLeuProSerGluMetSerProAspLeuHisArgLysValSerValGlnTyr 264
 Db 77 -----GATTTAATGAGAGAAATTCGACGCGAGAAAG 48
 Qy 265 CysAspValLeuHisLysLysAlaLeuArgAsnLysLys 277
 Db 47 ATGGAAGACACCAAAAAA-----9

RESULT 15

US-09-801-368-285
 Sequence 285, Application US/09801368
 Patent No. US20020128250A1
 GENERAL INFORMATION:
 APPLICANT: Busby, Robert
 APPLICANT: Cali, Brian
 APPLICANT: Hecht, Peter
 APPLICANT: Holtzman, Doug
 APPLICANT: Madden, Kevin
 APPLICANT: Maxon, Mary
 APPLICANT: Milne, Todd
 APPLICANT: Royer, John
 APPLICANT: No. US20020128250Alman, Thea
 APPLICANT: Salama, Sofie
 APPLICANT: Sherman, Amir
 APPLICANT: Silva, Jeff
 APPLICANT: Summers, Eric
 TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
 FILE REFERENCE: 109272.147
 CURRENT APPLICATION NUMBER: US/09/801,368
 PRIOR FILING DATE: 2001-03-07
 PRIOR APPLICATION NUMBER: US 09/487,558
 PRIOR FILING DATE: 2000-01-19
 PRIOR APPLICATION NUMBER: US 60/160,587
 NUMBER OF SEQ ID NOS: 440
 SOFTWARE: Patent In version 3.0
 SEQ ID NO 285
 LENGTH: 930
 TYPE: DNA
 ORGANISM: Saccharomyces cerevisiae
 US-09-801-368-285

Alignment Scores:
 Pred. No.: 8,69e-15 Length: 930
 Score: 264.50 Matches: 71
 Percent Similarity: 49.15% Conservative: 45
 Best Local Similarity: 30.08% Mismatches: 71
 Query Match: 15.28% Indels: 49
 DB: 10 Gaps: 5

US-09-709-103-3f1 (1-332) x US-09-801-368-285 (1-930)

Qy 76 TyrArgMetValIleLeuGlySerSerLysValGlyLysThrAlaIleValSerArgPhe 95
 Db 31 TATAAGATAGTAGTTCGCGTGGAGTGGCGTTGGTAAATCTGCTTAAACAATTCATTC 90
 Qy 96 LeuThrGlyArgPheGluAspAlaTyrThrProThrIleGluAspPheHisArgLysPhe 115

Db	91	ATTCAATCATACTTTGTGACGAATATGACCCCTACTATGGAAGATCTTTACGAAAAA	150
Qy	116	ThySerIleargGlyIValIYrGlnLeuAspIleLeuAspThrSerGlyAsnHisPro	135
Db	151	GTTGTGTCATCGATGACAAAGATATCCATTTTGGACATTCGTGATACCTGCGAAGAAG	210
Qy	136	PheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIleLeuValPheSer	155
Db	211	TATTCGTGGAGAGAGAAACAGTACATGAGACGTGGGGAAGGTTTCTTACTGGTCTATTC	270
Qy	211	LeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGlnIleLeuAspThr	175
Qy	156	LeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGlnIleLeuAspThr	175
Db	271	GTCACCTCTAGAAATCTCTTTGAGAGGTATCTGCTATATATACGCAAAAT	321
Qy	176	LysserCysLeuLysAsnLysThrLysGluAsnValAspValProLeuValIleCysGly	195
Db	322	CAAGAGATTAAGATCTCTACTACTACTCTGTAGTCGTGGTAGGT	366
Qy	196	AsnLysGlyAspArgAspPheTyrArgGluValAspGln	211
Db	367	AACCAATGAGACCTTGAAAAATGAAAGACAACTCTCTTATGAAAGCGGTACCCGTGCC	426
Qy	212	GlnGlnLeuValGlyAspAspProGlnArgCysAlaTyrPheGluIleSerAlaLys	231
Db	427	AAGCAGTTG	468
		-----AATGACACCTTTCTAGAAACGCTGGGAACAA	
Qy	232	AsnSerSerLeuAspGlnMetPheArgAlaLeuPheAlaMet	245
Db	469	GCCCATCAAGTAGACGAGCCCTTTATATACCTTATCTGTTGTAAGGACGACGGTGGG	528
Qy	246	-----AlaLysLeu	248
Db	529	AAATACATATGACATGAATGTCACATCGATTAATACGAATGAATTAAGATTCGGACCTA	588
Qy	249	ProSerGluMetSerProAspLeuHisArgLysValSerValGlnTyrCysAspValLeu	268
Db	589	ACCTCATCTGCAACACCGGATAGAGAAAAAAGAACACACGGGTCTTA	636
Qy	269	HisIleLysAlaLeuArgAsnLysLysLeuLeuArgAlaGlySerGly	284
Db	637	-----GTACTCGATATATCTTTGACCAATGCTGGACACTGGC	672

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 28, 2002, 01:18:22 ; Search time 173.866 Seconds
(without alignments)
10957.793 Million cell updates/sec

Title: US-09-709-103-1
Perfect score: 846
Sequence: 1 atgaactggcgcgatgat.....agcgtgcgcagtag 846

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002.*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	846	100.0	846	21	AAZ36892
2	846	100.0	1740	21	AAZ36914
3	846	100.0	1801	21	AAZ36893
4	846	100.0	1841	20	AAZ23022
5	844.4	99.8	1776	21	AAA49177
6	678	80.1	1689	20	AAZ23024
7	625	73.9	3986	20	AAZ23023
8	531	62.8	3079	20	AAZ23025
9	407.4	48.2	837	21	AAZ36913

1	846	100.0	846	21	AAZ36892
2	846	100.0	1740	21	AAZ36914
3	846	100.0	1801	21	AAZ36893
4	846	100.0	1841	20	AAZ23022
5	844.4	99.8	1776	21	AAA49177
6	678	80.1	1689	20	AAZ23024
7	625	73.9	3986	20	AAZ23023
8	531	62.8	3079	20	AAZ23025
9	407.4	48.2	837	21	AAZ36913

1	846	100.0	846	21	AAZ36892
2	846	100.0	1740	21	AAZ36914
3	846	100.0	1801	21	AAZ36893
4	846	100.0	1841	20	AAZ23022
5	844.4	99.8	1776	21	AAA49177
6	678	80.1	1689	20	AAZ23024
7	625	73.9	3986	20	AAZ23023
8	531	62.8	3079	20	AAZ23025
9	407.4	48.2	837	21	AAZ36913

10	407.4	48.2	2832	24	ABL92076	Human Tumour Endot
11	407.4	48.2	2973	24	ABL92087	Human Tumour Endot
12	407.4	48.2	3427	24	ABK71563	Human dithp polyu
13	358	42.3	3020	24	ABL92134	Mouse Tumour Endot
14	333.6	39.4	951	23	AA590571	DNA encoding novel
15	253.8	30.0	624	24	ABQ28320	Oligonucleotide fo
16	253.8	30.0	624	24	ABQ28321	Oligonucleotide fo
17	252.4	29.8	624	24	ABQ28318	Oligonucleotide fo
18	252.4	29.8	624	24	ABQ28319	Oligonucleotide fo
19	191.4	22.6	1305	23	ABL07789	Drosophila melanog
20	173.2	20.5	368	22	ABA51382	Human breast cell
21	173.2	20.5	368	22	ABA69388	Human foetal liver
22	173.2	20.5	368	22	ABA36323	Probe #14789 for g
23	173.2	20.5	368	22	AAK17661	Human brain expres
24	173.2	20.5	368	22	AAK3477	Human bone marrow
25	173.2	20.5	368	22	AAI24259	Probe #14192 for g
26	173.2	20.5	368	22	AAI49341	Probe #18227 used
27	173.2	20.5	368	22	AAI09818	Probe #9809 used t
28	173.2	20.5	368	22	AAI17600	Human genome-deriv
29	132	15.6	3061	24	ABQ72648	Human MDR1 encodin
30	125.8	14.9	1249	24	ABQ72648	Human MDR1 encodin
31	113.4	13.4	1144	22	AAH90075	Human bone marrow
32	113.4	13.4	4933	22	AAK81235	Human immune/haema
33	113.4	13.4	4934	22	AAK81234	Human polynucleoti
34	112.8	13.3	1087	22	AAI60838	Human polynucleoti
35	112.8	13.3	1108	22	AAI59052	Human intracellular
36	112.8	13.3	4167	24	AAAD37605	Human intracellular
37	107.2	12.7	558	11	AAQ03212	RAP2 Gene encoding
38	107.2	12.7	3300	22	AAI58569	Human polynucleoti
39	101	11.9	4543	23	ABL07788	Drosophila melanog
40	95.6	11.3	2223	22	AAH41597	Human H-Ras, c-Raf
41	95.6	11.3	2223	24	ABL50184	Human H-Ras, c-Raf
42	95.6	11.3	2223	24	ABL50216	Human H-Ras, c-Raf
43	95.6	11.3	2295	24	ABL50201	Human H-Ras, c-Raf
44	95.6	11.3	2295	24	ABL50233	Human H-Ras, c-Raf
45	94	11.1	2223	22	AAH41609	Human H-Ras, c-Raf

ALIGNMENTS

RESULT 1
AAZ36892
ID AAZ36892 standard; cDNA; 846 BP.
XX
AC AAZ36892;
XX
DT 13-MAR-2000 (first entry)
XX
DE CDNA encoding an activator of G protein signalling (AGS) protein.
XX
KW Activator of G protein signalling; AGS; ras-related G protein;
KW GTP hydrolysis; G protein activity; AGS; ras-related G protein;
KW G protein-coupled signal transduction; G-protein response pathway;
KW cellular signal transduction; ss.
XX
OS Homo sapiens.
FH Key Location/Qualifiers
CDS 1..846
FT /*tag= a
FT /product= "activator of G protein signalling (AGS) protein"
XX
PN CDNA encoding an a
XX DNA encoding an a
XX PD CDNA encoding an a
XX PD Human kd312 polype
XX PF CDNA encoding huma
XX PF Rat kd312 polypept
XX PR Human kd312 genomi
XX PR Rat kd312 genomic
XX PR DNA encoding a hom

(CADU-) CADUS PHARM CORP.

Cismowski M, Duzic E;
WPI; 2000-072337/06.
P-PSDB; AAY53921.

A new activator of G protein signalling used to treat disorders characterized by an aberrant AGS protein activity -

Claim 3; Fig 3A; 162pp; English.

The present sequence encodes an activator of G protein signalling (AGS) protein. The cDNA sequence was isolated from a human liver cDNA library. The AGS protein exhibits homology to ras-related G proteins, and contains alterations in conserved amino acids consistent with a deficiency in GTP hydrolysis activity. AGS stimulates G protein activity, G protein-coupled signal transduction and the pheromone response pathway in a receptor-independent manner. The AGS protein also shows G-gamma selectivity, as measured by growth assays in yeast expressing various mammalian G-gamma constructs, and tissue-specific expression, as measured by Northern blot analysis. The AGS protein can be used to screen for compounds that modulate cellular signal transduction. The protein is used to treat disorders characterized by an aberrant AGS protein activity or AGS nucleic acid expression.

Sequence 846 BP; 176 A; 288 C; 252 G; 130 T; 0 other;

Query Match	100.0%;	Score 846;	DB 21;	Length 846;
Best Local Similarity	100.0%;	Pred. No. 1.7e-141;		
Matches 846;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

```

QY      1  ATGAAACTGGCGCGATGATCAAGAAGATGTGCCGAGCGACTCGGAGCTGAGTATCCCG 60
Db      |||
QY      1  ATGAAACTGGCGCGATGATCAAGAAGATGTGCCGAGCGACTCGGAGCTGAGTATCCCG 60
Db      |||

QY      61  GCCAAGAACTGATTCGATGGTTCATCCTCGGCTCTGCCAAGTGGCGAAGACGGCCATC 120
Db      |||

QY      61  GCCAAGAACTGATTCGATGGTTCATCCTCGGCTCTGCCAAGTGGCGAAGACGGCCATC 120
Db      |||

QY      121  GTGTCGCGGTTCTCCACGGCGCGTTTCGAGAGCGCTACACGCTACCATCGAGCACTTC 180
Db      |||

QY      121  GTGTCGCGGTTCTCCACGGCGCGTTTCGAGAGCGCTACACGCTACCATCGAGCACTTC 180
Db      |||

QY      181  CACCGCAAGTCTTACTCCATCGCGGGGAGGTCTACAGCTCGACATCTCGACACGTCC 240
Db      |||

QY      181  CACCGCAAGTCTTACTCCATCGCGGGGAGGTCTACAGCTCGACATCTCGACACGTCC 240
Db      |||

QY      241  GGCAACACCCGTTCCCGCCCATGCGGCGCTCTCCATCTCAGAGGACGTTTTCATC 300
Db      |||

QY      241  GGCAACACCCGTTCCCGCCCATGCGGCGCTCTCCATCTCAGAGGACGTTTTCATC 300
Db      |||

QY      301  CTGGTGTTCAGTCTGACAAACCGGACACTCTTCGAGGAGGTGCAGCGCTCAGCGCAGCAG 360
Db      |||

QY      301  CTGGTGTTCAGTCTGACAAACCGGACACTCTTCGAGGAGGTGCAGCGCTCAGCGCAGCAG 360
Db      |||

QY      361  ATCTCTGACACCAAGTCTTCGCTCAAGAACAAACCAAGGAGACGTGGAGCTGCCCTG 420
Db      |||

QY      361  ATCTCTGACACCAAGTCTTCGCTCAAGAACAAACCAAGGAGACGTGGAGCTGCCCTG 420
Db      |||

QY      421  GTCATCTGGCGGAACAAGGTTGACCGGACACTCTTACCGGAGGTGGACCAAGCGCGATC 480
Db      |||

QY      421  GTCATCTGGCGGAACAAGGTTGACCGGACACTCTTACCGGAGGTGGACCAAGCGCGATC 480
Db      |||

QY      481  GAGCAGCTGGTGGCGGACGACGCCAGCGCTCGCGCTACTTCGAGATCTCGGCCAAGAAG 540
Db      |||

QY      481  GAGCAGCTGGTGGCGGACGACGCCAGCGCTCGCGCTACTTCGAGATCTCGGCCAAGAAG 540
Db      |||

QY      541  AACAGCAGCTGGACACAGATGTTCCCGCGCTCTTCGCCATGGCCCAAGCTCCCGCAGCAG 600
Db      |||

QY      541  AACAGCAGCTGGACACAGATGTTCCCGCGCTCTTCGCCATGGCCCAAGCTCCCGCAGCAG 600
Db      |||

```

yeast expressing various mammalian G-gamma constructs, and tissue-specific expression, as measured by Northern blot analysis. The AGS protein can be used to screen for compounds that modulate cellular signal transduction. The protein is used to treat disorders characterized by an aberrant AGS protein activity or AGS nucleic acid expression.

Mon Dec 30 09:16:48 2002

```

QY 841 AGCTAG 846
Db 1020 AGCTAG 1025

RESULT 6
AAZ23024
ID AAZ23024 standard; cDNA; 1689 BP.
AC AAZ23024;
DT 17-JAN-2000 (first entry)
DE Rat kd312 polypeptide encoding cDNA.
XX kd312 polypeptide; cancer; immunodeficiency disorder; AIDS; stroke;
XX heart attack; head trauma; neurodegenerative disease; rat;
KW Parkinson's disease; Alzheimer's disease; ss.
XX Rattus sp.
OS W09950288-A2.
PN 07-OCT-1999.
XX 30-MAR-1999; 99WO-US06993.
XX 31-MAR-1998; 98US-0053374.
XX (AMGE-) AMGEN INC.
XX YEN K;
XX WPI: 1999-601322/51.
XX P-PSDB; ANY42694.
XX kd312 polypeptides useful for treating diseases and disorders
XX associated with alterations in cell proliferation and cell death -
XX Claim 2; Fig 9; 85pp; English.
XX The invention provides nucleic acid molecules encoding human and rat
XX kd312 polypeptides. The kd312 polypeptides can be expressed by standard
XX recombinant methodology. The kd312 sequences, and the antibodies against
XX the proteins may be used to treat or diagnose the presence or progression
XX of various diseases, e.g. cancer, immunodeficiency disorders (e.g. AIDS),
XX stroke, heart attack, head trauma, and neurodegenerative diseases (e.g.
XX Parkinson's disease and Alzheimer's disease). The present sequence
XX represents the rat kd312 cDNA sequence.
XX
XX Sequence 1689 BP; 417 A; 462 C; 434 G; 376 T; 0 other;
XX
Query Match 80.18; Score 678; DB 20; Length 1689;
Best Local Similarity 88.4%; Pred. No. 1.1e-111; Indels 3; Gaps 1;
Matches 748; Conservative 0; Mismatches 95;
QY 1 ATCAACTGGCGCGATGATCAAGAAGATGTCCCGGAGCGACTCGGAGCTGAGTATCCCG 60
Db 132 ATGAAGTGGCGCGATGATCAAGAAGATGTCCCGGAGCGACTCGGAGCTGAGTATCCCG 191
QY 61 GCCAAGAACTGATATCGAATGATGTCCTCGGCTGTCCTCAAGGTGGGCAAGCGGCATC 120
Db 192 GCCAAGAACTGATACAGGATGTCCTCGGCTGTCCTCAAGGTGGGCAAGCGGCATC 251
QY 121 GTGTCGGCTTCTCAACCGCGCTTCGAGGAGCGCTACAGCCTACACCTACGAGACTTC 180
Db 252 GTGTCGGCTTCTCAACCGCGCTTCGAGGAGCGCTTACACCTACACCTACGAGACTTC 311
QY 181 CACCGCAAGTTCTACCTACCGCGCGAGTCTACAGCTCGACATCTCGACACGTC 240
Db 312 CACCGCAAGTTTCTACCTACCGCGCGAGTCTACAGCTCGACATCTCGACACATCT 371

```

```

QY 241 GGCAACACCCGTTCCCGCATGCGCGCTCTCATCTCCACAGGAGACGCTTTTCATC 300
Db 372 GGCAATCATCCGTTTCCCGCATGCGCGCTCTCTATCTCCACAGGAGACGCTTTTCAT 431
QY 301 CTGCTGTTCACTTGGACAAACCGGACTCTTTCGAGGAGGTGCAGCGCTCAGGACGAG 360
Db 432 CTGCTGTTCACTTAGACAAACCGGACTCTTTCGAGGAGGTGCAGGAGGTCAACAGCAG 491
QY 361 ATCTCGACACCAAGTCTTGCCTCAAGAAACAAACAAAGGAGAACTGGACGCTGCCCCG 420
Db 492 ATCTTAGACACCAAGTCTTGCCTCAAGAAACAAACAAAGGAGAACTGGACGCTGCCCCG 551
QY 421 GTCATCTGCGCAACAAAGGAGTGCAGGACTTCTACCGCAAGTGGAGCAGCGGAGATT 480
Db 552 GTCATTTGCGGTAAACAAAGGAGTGCAGGACTTCTACCGCAAGTGGAGCAGCGGAGATT 611
QY 481 GAGCAGCTGGTGGGAGGAGCGGAGCGGCTGCGCTTCTCGAGATCTCGGCAAGAAAG 540
Db 612 GAGCAGCTGGTGGGAGGAGCGGCTGCGCTTCTCGAGATCTCGGCAAGAAAG 671
QY 541 AACAGCAGCTGGGAGGAGTGTTCGCGGCTTTCGCGCAAGTGGGCAAGCTGCCAGCGGAG 600
Db 672 AATAGCAGCTGGGAGGAGTGTTCGCGGCTTTCGCGCAAGTGGGCAAGCTGCCAGCGGAG 731
QY 601 ATGAGCCAGACTGCGGAGGAGCGGCTGCGCTTCTCGAGATCTCGGCAAGAAAG 660
Db 732 ATGAGCCAGACTGCGGAGGAGCGGCTGCGCTTCTCGAGATCTCGGCAAGAAAG 791
QY 661 GCGCTGCGCAACAAAGGAGTGTTCGCGGCTTTCGCGCAAGTGGGCAAGCTGCCAGCGGAG 720
Db 792 GCTCTGAGGAGCAAGGAGTGTTCGCGGCTTTCGCGCAAGTGGGCAAGCTGCCAGCGGAG 848
QY 721 GAGCCTTTGGCATGCTGCGGAGCGGCTTTCGCGGCTTTCGCGCAAGTGGGCAAGCTGCCAG 780
Db 849 GATGCTTTGGCATGCTGCGGAGCGGCTTTCGCGGCTTTCGCGCAAGTGGGCAAGCTGCCAG 908
QY 781 TACATCGCGCAAGGAGGAGTGTTCGCGGCTTTCGCGGCTTTCGCGCAAGTGGGCAAGCTGCCAG 840
Db 909 TACATCGCGCAAGGAGGAGTGTTCGCGGCTTTCGCGGCTTTCGCGCAAGTGGGCAAGCTGCCAG 968
QY 841 AGCTAG 846
Db 969 AGTTAG 974

RESULT 7
AAZ23023
ID AAZ23023 standard; DNA; 3986 BP.
XX
XX AAZ23023;
XX 17-JAN-2000 (first entry)
XX Human kd312 genomic DNA sequence.
XX kd312 polypeptide; cancer; immunodeficiency disorder; AIDS; stroke;
XX heart attack; head trauma; neurodegenerative disease; human;
XX Parkinson's disease; Alzheimer's disease; ss.
XX Homo sapiens.
XX W09950288-A2.
XX 07-OCT-1999.
XX 30-MAR-1999; 99WO-US06993.
XX 31-MAR-1998; 98US-0053374.
XX (AMGE-) AMGEN INC.
XX YEN K;
XX

```


QY	1	ATGAACTGGCCGGATGATCAAGAGATGTCCGGAGGACTCGAGACTGAGTATCCCG	60
DB	303	ATGAACTGGCCGGATGATCAAGAGATGTCCCGAAGGACTCTGAACTGAGTATCCCG	362
QY	61	GCCAAAGTCTATCGCATGTGTATCCTCGGCTCGTCCAAAGTGGGCAAGAGCGCATC	120
DB	363	GCCAAAGTCTATCGCATGTGTATCCTCGGCTCATCAAAGTGGGCAAGAGCGCATC	422
QY	121	GTGTGGCGCTTCTACCGCGCCTTCGAGGAGCGCTACAGCTTACCATCGAGACTTC	180
DB	423	GTGTGGCGCTTCTACCGCGCGCTTCGAGGAGCGCTTACAGCTTACCATCGAGACTTC	482
QY	181	CACCCAAAGTTTACTCATCTCCGCGGAGGTTCTACAGCTCGACATCCTCGACACGTCC	240
DB	483	CACCCAAAGTTTACTCATCTCCGCGGGAAGTCTACCAAGTGGACATCTGGACACATCT	542
QY	241	GGCAACACCGCTTCCCGCGCATCGCGCGCTCTCATCTCTCAC-	284
DB	543	GGCAATCATCGCTTCCCGCATCGCGCGCTCTCATCTCACAGGTGAGTGGGGACC	602
QY	285	-----	284
DB	603	GACAGGACCGTGGGGAGGAACTCGCGGGAGCGGATGGGCGGTGTGTGCTTGGG	662
QY	285	-----	284
DB	663	GCTGTGCTGTCTGTCTCCCTGTGTGGCAGCTGCCCTCACTTTCCACTCGTTCCCTTG	722
QY	285	-AGGAGACGTTTTCATCTGTGTTCAGTCTGGACAAACCGGACTCTCTTCGAGGAGTGC	343
DB	723	TAGGAGACGTTTTCATCTGTGTTCAGCTTAGCAACCGGACTCTCTTCGAGGAGTGC	782
QY	344	AGCGGCTCAGGACGAGATCTCTGCACACCAAGTCTTGCTCTCAAGACAAACCAAGGAGA	403
DB	783	AAAGGCTCAACAGCAGATCTCTAGACACCAAGTCTGTCTCAAGACAAACCAAGGAGA	842
QY	404	ACGTGACCTGCCCTTGCTATCTGCGCAACAGGCTGACCGGACTTCTTACCGCGAGG	463
DB	843	ATGTGGACGTGGCGTGTGTCTATTTGCGGTACAAAGGGAGCCGGACTTCTACCGCGAAG	902
QY	464	TGGACACCGCGAGATPCGAGCAGTGTGTGGCGAGCAGCCCGAGCGCTGGCCTACTTCG	523
DB	903	TGGACACCGGGAGATTGAGCAGCTGTGTGGCGAGTACCCCTCAGCGCTTGTGCTACTTCG	962
QY	524	AGATCTGGCAGAGAGAACACACCTGGACACAGATGTTCGCGCGGCTCTTCGCGCATGG	583
DB	963	AGATCTGGCAGAGAGAAATACACCTTGGACAGAGTTCCTGCGCTCTTTGGCCATGG	1022
QY	584	CCAAAGCTGCCACCGAGATGAGCCAGACCTGCACCGGAGGCTCGGTGCATCTGGG	643
DB	1023	CCAAAGCTGCCACCGAGATGAGCCCTGTGCTGACCCGAAAGGTGTCTGTGCATGTACTGTG	1082
QY	644	ACGTGTGCACAGAGAGCGCTGCGGAAACAGAGAGCTGTCTGCGGGCGCGCAGCGCGGG	703
DB	1083	ACGTGTGCACAGAGAGAGCTCTGAGGACAAAGAGCTTCTGCGTGCGGGCA---GCGGAG	1139
QY	704	CGCGCGCGACCCGGCGAGCGCTTTGGGATCGTGGCAACCTTTCGCGCGCGCGCCAGCG	763
DB	1140	GTGGGGCGCACCGGAGATGCTTTGGCATCTTTGGCGCCCTTTGCTTCGAGACCTTAGCG	1199
QY	764	TACACAGGACCTCATGTACATCCGGAAGAGCGCGCGGACGCGAGCGACGAGCCAAAGGACA	823
DB	1200	TGCATAGGACCTCATGTACATCTGTGAGAAACCAAGTGTGACAGCCAGGCTAAGGACA	1259
QY	824	AGGAGCGCTGCGTCAATCAGCTAG	846
DB	1260	AGGAGCGCTGTCTCAGTTAG	1282

RESULT 9
AAZ36913
TD AAZ36913 standard: DNA: 837 BP.

XX	AAZ36913;
XX	13-MAR-2000 (first entry)
XX	DNA encoding a homologue of G protein signalling AGS1.
XX	Activator of G protein signalling; AGS; ras-related G protein;
XX	GTP hydrolysis; G protein activity; pheromone response pathway;
XX	G protein-coupled signal transduction; G-gamma selectivity;
XX	cellular signal transduction; AGS1 homologue; ss.
OS	Homo sapiens.
XX	Key Location/Qualifiers
FT	1..837
FT	/*tag= a
FT	/product= "AGS1 homologue"
XX	W09958670-A1.
PN	18-NOV-1999.
PD	07-MAY-1999; 99WO-US10151.
XX	08-MAY-1998; 98US-0084842.
XX	07-OCT-1998; 98US-0103955.
XX	(CADU-) CADUS PHARM CORP.
XX	Cismowski M, Duzic E;
XX	WPI; 2000-072337/06.
XX	P-PSDB; AAY53923.
XX	A new activator of G protein signalling used to treat disorders
PT	characterized by an aberrant AGS protein activity -
XX	Example 15; Page 143-144; 162pp; English.
XX	The present sequence encodes a homologue of a human AGS1 (activator
CC	of G protein signalling (AGS)) protein. The AGS cDNA sequence was
CC	isolated from a human liver cDNA library. The AGS protein exhibits
CC	homology to ras-related G proteins, and contains alterations in
CC	conserved amino acids consistent with a deficiency in GTP hydrolysis
CC	activity. AGS stimulates G protein activity, G protein-coupled signal
CC	transduction and the pheromone response pathway in a receptor-independent
CC	manner. The AGS protein also shows G-gamma selectivity, as measured by
CC	growth assays in yeast expressing various mammalian G-gamma constructs,
CC	and tissue-specific expression, as measured by Northern blot analysis.
CC	The AGS protein can be used to screen for compounds that modulate
CC	cellular signal transduction. The protein is used to treat disorders
CC	characterized by an aberrant AGS protein activity or AGS nucleic acid
CC	expression.
XX	sequence 837 BP; 177 A; 272 C; 233 G; 155 T; 0 other;

Query Match 48.2%; Score 407.4; DB 21; Length 837;
Best Local Similarity 71.5%; Pred. No. 11e-63;
Matches 594; Conservative 0; Mismatches 201; Indels 36; Gaps 3;

QY	193	TACTCATCGGGGAGGTCTACAGCTCGACATCTCGACAGTCCGCAACACCG	252
Db	214	TACAACATCGGGGACATGTACCAGCTCGACATCTGATACCTCTGCAACACCGC	273
QY	253	TTCCCGCCATCGGGGCTCTCCATCTCACAGAGACGTTTTCATCTCTGGTTCAGT	312
Db	274	TTCCCGCCATCGCAGGCTGTCCATCTCACAGGATGCTTTCATCTCTGGTGTTCAG	333
QY	313	CTGGACACCGGACTCTTTCAGAGGTTGAGGGCTCAGGCAGCAGATCTCGACACC	372
Db	334	CTGGATACCGGAGTCTTCGATGAGGTCAAGGCTTCAGAAGCAGATCTCGGAGTC	393
QY	373	AGTCTTGCTTAAAGACAAACCAAGGAAAGTGCAGTGCCTCTGTCATCTCGGC	432
Db	394	AGTCTTGCTTGAAGACAAAGACCAAGGAGCGCGGAGCTGCCATGTCATCTGGC	453
QY	433	AAACAAGGTGACC---GCGACTTCTACCGAGGTGGACAGCGGAGATCGACAGCTG	489
Db	454	AAACAAGACGACACGCGGAGCTGTGCGCGCAGGTGCCACACCGAGCGCGAGTGTG	513
QY	490	GTGGCGCAGACCCCAAGCGCTGCGCTACTTGGAGATCTGGCCAAAGAACAGCAGC	549
Db	514	GTGTCGGCGAGC---AGAACTCGCGCTACTTCGAGTGTGCGCCAAAGAACACAC	570
QY	550	CTGACACAGATGTTTCGCGCGCTTTCGCCATGGCCAGCTGCCAGCGAGATGAGCCCA	609
Db	571	GTGACAGAGATGTTCTACGTGCTCTTCAGATGGCCAGCTGCCACGAGATGAGCCCC	630
QY	610	GACCTCACCGCAAGTCTCGGTGAGTACTCGGACGTGCTGCAACAAGAGCGCTCGG	669
Db	631	GCCCTGATCGCAAGATCTCCGTGCAAGTACGTGTGACGCTTCCACCCCAAGGCC	683
QY	670	AAACAAGTGTGTGCGGCGCGCACGCGCGCGCGCGCGCGCGCGAGCGCTTT	729
Db	684	-----CTTCTGCATGCGCCCGCTCAAGGATGGAGCGCTAT	720
QY	730	GGCATCGTGGACCCCTTCGCGCGCGCGCCAGCTTACACGCGACCTCATGTACATCCGC	789
Db	721	GGCATGTCTCGCCCTTCGCGCGCGCGCCAGCTTACAGTGAACCTTCAAGTACATCAAG	780
QY	790	GAGAGCGCAGCGCGGACCGACCGCCAAAGCAAGGAGCGCTGCGTCAATC	840
Db	781	GCCAAGTCTTGGGAGGCGCAGCGCGTTCAGAGGAGCAAGTGCACCATC	831

RESULT	10
ABL92076	
ID	ABL92076 standard; cDNA; 2832 BP.
XX	
AC	ABL92076;
XX	
DT	30-MAY-2002 (first entry)
XX	
DE	Human Tumour Endothelial Marker polynucleotide SEQ ID NO 174.
KW	Human; mouse; rat; TEM; tumour endothelial marker; NEM; PEM; cytostatic normal endothelial marker; pan-endothelial marker; immunostimulant;
KW	antiangiogenic; tumour; neovascularisation; vasculatured tumour;
KW	papillary kidney disease; diabetes; retinopathy; rheumatoid arthritis;
KW	psoriasis; gene; ss.
XX	
OS	Homo sapiens.
XX	
PEN	WO200210217-A2.
XX	
PD	07-FEB-2002.
XX	
PF	01-AUG-2001; 2001WO-US24031.
XX	
RPR	02-AUG-2000; 2000US-222599P.
RPR	11-AUG-2000; 2000US-224360P.
RPR	11-APR-2001; 2001US-282850P.
XX	

PA (UYJO) UNIV JOHNS HOPKINS.
PX
PI St Crolx B, Kinzler KW, Vogelstein B;
XX
XX WPI; 2002-291856/33.
DR
XX
PX An isolated molecule comprising an antibody variable region which
PT specifically binds to an extracellular domain of a tumor endothelial
PT marker (TEM) protein, useful for inhibiting tumor growth -
XX
XX Disclosure; Page 119-120; 331pp; English.
XX
XX The invention relates to an isolated molecule comprising an antibody
CC variable region which specifically binds to an extracellular domain of a
CC tumor endothelial marker (TEM) protein selected from ABB90732, ABB90740,
CC ABB90749, ABB90750 and ABB90769. The antibodies which bind to TEM
CC proteins have cytostatic, immunostimulant and angiogenic activity.
CC They are useful for inhibiting tumor growth, neoangiogenesis in
CC subjects bearing a vascularized tumor, polycystic kidney disease,
CC diabetic retinopathy, rheumatoid arthritis and psoriasis. Human, mouse
CC and rat TEM genes and the encoded proteins (ABL92075-ABL92141 and
CC ABB90721-ABB90789) are disclosed, as are marker oligonucleotide
CC sequences: tumor endothelial markers (TEM) ABL91936-ABL92041 and
CC ABL92143-ABL92191; normal endothelial markers (NEM) ABL92042-ABL92074;
CC and pan-endothelial markers (PEM) ABL91903-ABL91995.
XX
XX Sequence 2832 BP; 602 A; 891 C; 727 G; 612 T; 0 other;
SQ

Query Match	48.2%;	Score 407.4;	DB 24;	Length 2832;
Best Local Similarity	71.5%;	pred. No. 11e-63;		
Matches 594;	Conservative 0;	Mismatches 201;	Indels 36;	Gaps 3;
QY	13	GGGATGATCAAGAGATGTGCCGAGCGATCTGGAGCTCAGTATCCCGCCGAAGAACTGC	72	
Db	34	GCATGATGAGACTTTGTCCAGCGGAAGTCAGCCTCAGTGTGCCGCCCAAAACTCA	93	
QY	73	TATCGCATGGTCATCCTCGGCTCGTCCAAGTGGGCAAGCGCCATCGTGTCCGGCTTC	132	
Db	94	TACCGCATGTGTGTGGTGGCTCTCGGGTGGGCAAGAGCTCCATCGTGTCTGGCTTC	153	
QY	133	CTCACGGCGCCTTCGAGGAGCGCTACACGCCCTACCATCGAGGACTTCCACCGCAAGTTC	192	
Db	154	CTCAATGCCGCTTTGAGGACCACTACACCCACCATCGAGGACTTCCACCGTAAGSTA	213	
QY	193	TACTCCATCCGCGCGGAGGTTCTACCAAGCTCGACATCCTCGACAGCTCCGGCAACCCACCG	252	
Db	214	TACAACATCCGCGCGACATGTACCAAGCTCGACATCCTGGATACCTCTGGCAACACCCCC	273	
QY	253	TTCCCCGCCATGGCGGCTCTCCATCTCACAGGAGCGTTTTCATCTCTGGTGTTCAGT	312	
Db	274	TTCCCCGCCATGGCAGGCTGTCCATCTCACAGGGGATGTCTTCATCTCTGGTGTTCAGC	333	
QY	313	CTGGACMACCGCGACTCTCTCGAGGAGGTGACAGGGCTCAGGACGAGAGTCTCGACAC	372	
Db	334	CTGGATAACCGGAGTCTTCGATGAGGTCAAGCCCTTCAGAGAGAGATCTCTGGAGGTC	393	
QY	373	AAGTCTTGCTCAAGAACAAACCAAGAGAGAACGTGGACGTGCCCTCTGTCTCGCGC	432	
Db	394	AAGTCTTGCTGAGNACAAAGACCAAGAGGCGCGGAGTGTGCCATGTGTCTCTGTGGC	453	
QY	433	ACAAGGGTGAC---GGAGCTTCTACCGCGAGGTGACACAGCGCGAGATCGACAGCTG	489	
Db	454	AACAAGACGACCGGGGAGTGTGGCGCGAGTGTGCCACCCAGGCGCGAGTGTGTG	513	
QY	490	GTGGCGACGACCCCGAGCTGGCCCTACTTCAGATCTCGGCCAAGAGAACAGCAGC	549	
Db	514	GTGTCGGCGAGC---AGACTGGCGCTACTTGGAGTGTGGCCCAAGAGAACACCAAC	570	
QY	550	CTGGACCAAGATGTCGCGCGCTCTTGGCATGTGACGCTACTGGACGTGCTGCACAAAGCGCTGCGG	609	
Db	571	GTGGACGAGATGTTACGTGCTCTTACGATGCCCCAAGCTGCCCCAGGATGAGCCCA	630	
QY	610	GACCTGCACCGCAAGTCTCGGTGCATCTACTGGACGTGCTGCACAAAGCGCTGCGG	669	

	Query Match	48.2%	Score 407.4	DB 24	Length 2973	
	Best Local Similarity	71.5%	Pred. No. 1.e-63			
	Matches 594	Conservative	0	Mismatches 201	Indels 36	Gaps
QY	13	GCATGATCAAGAAGATGTCGCCGAGCAGCTCGGAGCTGAGTATCCCGGGCCAAAGACTGC	72			
Db	205	GCATGATGAAGACTTTGTCCAGCGGGAACCTGCACGCTCAGTGTGCCGCCCAAAATCA	264			
QY	73	TATCGCATGGTCATCTTCGGGCTCGTCCAAAGTGGGCAAGACGGCCATCGTCTCGCGCTTC	132			
Db	265	TACCGCATGGTGTGTGGTGGCTCTCGGGTGGCAAGAGCTCCATCGTCTCGCGCTTC	324			
QY	133	CTCACCGCGCGCTTCGAGGACGCCTACAGCCTACCATCGAGGACTTCCACCGCAAGTTC	192			
Db	325	CTCAATGGCGGCTTTTGAGAGCACCAGTACACCCACCATCGAGGACTTCCACGGTANGSTA	384			

1

b		385	TACAACATCCGGCGGAGCATCTACCAGTGCACATCCTTGATACCTCTGTGCAACATCTTTGCAATCTTCCTGCTGCTGTTAGT	312
y		253	TTCCCGGCATCGCGCGCTCTCCATCCTCACAGGAGACGTTTTTCATCTCTGCTGTTGTTT	
b		445	TTCCCCGCCATCGCAGAGCTGTCCATCCTCACAGGGATGTCTTTCATCTCGTGTGTTCAGC	504
y		313	CTGGCAAACCGGACTCCTTCGAGAGGTGACGGCTCAGGCAGCAGATCCTCGACACC	372
b		505	CTGNTAACCGGAGTCTTCATGAGTCAAGCGCTTCAGAAGCAGATCCTGGAGGTC	564
y		373	AAGTCTTGCTCANACAAAACCAAGAGAACTGTGACGTGCCTCTGGTCACTCTCGCGC	432
b		565	AAGTCTTGCTGAAGAACAGACCAAGAGGCGGGAGGTGCCCATGGTCACTCTGTGC	624
y		433	ANCAAGSGTAGCC-- --GCGACTTCTACCGCGAGGTGGACACGCGGAGATCGAGCAGCTG	489
b		625	ANCAAGACGACACCGCGCAGCTGTGCCGCGAGTGTCCCACCGAGGCGCAGCTGCTG	684
y		490	GTTGGCGCAGACCCCCAGCGCTGGCGCTACTTTCAGATCTCGGCCAAGAAAGACACGACG	549
b		685	GTTGTCGGCGACG-- --AGAACTCGCGCTACTTTCGAGGTGTCTGGGCCAAGAAAGACCCNAC	741
y		550	CTGGACAGATGTTCCGCGCGCTTTCGCGCATGGCCAAGCTGCCACGAGATGAGCCCA	609
b		742	GTTGACAGAGATTCTACGTGCTTTTCACCATGGCCAAGTGCACACGAGATGAGCCCC	801
y		610	GACCTGCACCGCAAGSTTCTCGGTGCAGTAATACTCGCAGCTGCTGCACAGAGGCGCTCGGG	668

802	GCCCTGATCGACAGATCTCCGGTGCCTC	725
QY		
670	AACAAGAAGCTGTGCGGGCGGACGGGGGGGGGGGGCGGCCCGGCGAGCGCCTTT	725
D		
855	-----CTTCTGCATGCGCGCGTGCAAGGATGGAGCGCCTAT	891
D		
730	GGCATGTGGCACCCCTTCGCGCGCGGCCCGGTCACAGCGACTCATGTACATCCGC	789
QY		
892	GGCATGTGTCTCGCCCTTTTCGCCGCGCGCCCGCCCGCTCAACGTGACTCAAGTACATCAAG	951
D		
790	GGAAGGTCAGTCGCGCGCAGCCAGCGGCCAAGGACAAAGGCGCTCGGTGATC	840
:		

[illegible]

XX Human; dithp; diagnostic and therapeutic polynucleotide; gene; ss;
KW cell proliferative disorder; cancer; tumour; autoimmune disorder; bone
KW cell proliferative disorder; cancer; tumour; autoimmune disorder; bone

polycystic kidney disease; diabetes; retinopathy; rheumatoid arthritis; psoriasis; gene; ss.

Mus musculus.

WO200210217-A2.

07-FEB-2002.

01-AUG-2001; 2001WO-US24031.

02-AUG-2000; 2000US-222599P.

11-AUG-2000; 2000US-224360P.

11-APR-2001; 2001US-282850P.

(UYJO) UNIV JOHNS HOPKINS.

St Croix B, Kinzler KW, Vogelstein B;

WPI; 2002-291856/33.

P-PSDB; ABB90781.

An isolated molecule comprising an antibody variable region which specifically binds to an extracellular domain of a tumor endothelial marker (TEM) protein, useful for inhibiting tumor growth -

Disclosure; Page 294-295; 331pp; English.

The invention relates to an isolated molecule comprising an antibody variable region which specifically binds to an extracellular domain of a tumor endothelial marker (TEM) protein selected from ABB90732, ABB90740, ABB90749, ABB90750 and ABB90769. The antibodies which bind to TEM proteins have cytostatic, immunostimulant and antiangiogenic activity. They are useful for inhibiting tumor growth, polycystic kidney disease, subjects bearing a vascularised tumour, rheumatoid arthritis and psoriasis. Human, mouse and rat TEM genes and the encoded proteins (ABL92075-ABL92141 and ABB90721-ABB90789) are disclosed, as are marker oligonucleotide sequences: tumour endothelial markers (TEM) ABL91996-ABL92041 and ABL92143-ABL92191; normal endothelial markers (NEM) ABL92042-ABL92074; and pan-endothelial markers (PEM) ABL91903-ABL91995.

Sequence 3020 BP; 690 A; 848 C; 780 G; 702 T; 0 other;

Query Match 42.3%; Score 358; DB 24; Length 3020;

Best Local Similarity 66.2%; Pred. No. 6.2e-55;

Matches 548; Conservative 0; Mismatches 250; Indels 30; Gaps 1;

13 GCGATGATCAAGAGATGTCGCCGAGGACTCGGACTGAGTATCCCGCCAAAGACTGC 72
 373 GCCATGATGAAGACCTTGTCCAGTGGGAACTGCACACTCAATGTGCTCTAAGAACTCC 432
 73 TATCGATGTCATCTCGCTCGCTCAAGTGGGCAAGACGCCCATCTGTCTCGGCTTC 132
 433 TACCGATGTTGTTGCTGCTGCTCCGAGTGGGCAAGAGCTCCATCTCTCCGCTTC 492
 133 CTCACCGCCGCTTCGAGGAGCGCTACACCGCTACCATGAGGACTTCACCGCAGTTTC 192
 493 CTCATGCGCCCTTTGAGGACCACTACACGCCCATCTATCGAGGACTTTCATCGAAGTTG 552
 193 TACTTCATCCCGCGGAGGTCTACACGCTCGACATCTTCGACAGCTCCGCGCAACACCGC 252
 553 TACAACATCCACGGGAGATGATACGCTGATATCTTCGACACCTTGGCAACACCCCA 612
 253 TTCGCCGCTACCGGCGCTCTCCATCTCCTACAGGAGACGCTTTTCATCTGTTGTTCA 312
 613 TTCCTGCTGCTGCGCGCTCTCCATCTCTACAGGAGATGCTTTCATCTGTTGTTCA 672
 313 CTGGACACCGGAGCTCTTCGAGGAGGTGCGCGGCTCAGGAGGAGCTTCCTCGACACC 372
 673 CTGGATAGCGGGAGTCTCTTTGATGAGTCAAGCGCTTCAGAAACAGATCTTGGAGTC 732
 373 AAGTCTTCTCAAGAACAAACCAAGGAGAACCTGGACGTGCGCCCTGGTCTATCTGCGGC 432

733 AAGTCTCTGCTGAAGATAAACAAGAGGAGGAGGAGAGAGCTGCCCATGTGTGCTGGG 792
 433 AACAGGAGTACCGCGACTTCTACCGGAGGTGGACCGAGGATCGAGCAGCTGTTG 492
 793 AACAGAAATGACACACAGTGAAGTGTGCCCGAGTCCCTGCCATGGAGGCTGAGCTG 852
 493 GGCAGACACCCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 552
 853 GTGCTGCTGATGAAGAACTGCGCTATTTTCAGGTTGTACGCCAAGAAGAACATAT 912
 553 GACAGATGTTCCCGCGCTTTCGCGCATGCGCAAGCTGCCAGCGAGATGAGCCAGAC 612
 913 AACGAGATGTTCTATGTGCTTTCAGCATGCGCAAGCTGCCCGCATGAGATGAGCC 972
 613 CTGACACCGCAAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 672
 973 CTGACACATAGATCTCGTGCAGTGGCGGATGCTTTTTCACCCCGCGC----- 1022
 673 AGAAGTGTGCTGCG 732
 1023 -----CTTCTGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1062
 733 ATCTGTGCGACCTTTCG 792
 1063 ATGTTCTCACCTTTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1122
 793 AAGGCGAGCG 840
 1123 AAGTCTTACGGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1170

RESULT 14

AAS90571
 ID AAS90571 standard; cDNA; 951 BP.

XX AAS90571;

XX AAS90571;

DT 13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #26375.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 XX food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.

OS WO200175067-A2.

PN 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US08631.

XX 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX P-PSDB; ABB26384.

XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 XX Claim 1; SEQ ID No 26375; 103pp; English.

PS The invention relates to isolated polynucleotide (I) and

XX polypeptide (II) sequences. (I) is useful as hybridisation probes,

CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome

Mon Dec 30 09:16:48 2002

```

QY 501 CCCCCAGCGCTGCGCTTCTGAGATCTCGGCCAAGAAAGACAGCAGCCTGGACCAGAT 560
Db 560 CCCCCAGCGCTTACGCTTCTGAAATCTCGACCAAAAAAACAACACCTAAACCAAT 501
QY 561 GTTCCGCGCGCTCTTCCCATGCGCAAGCTGCCAGCGGAGATGAGCCAGACCTGCACCG 620
Db 500 ATTCCGCGCGCTCTTCCCATGCGCAAGCTGCCAGCGGAGATGAGCCAGACCTGCACCG 441
QY 621 CAAGGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 680
Db 440 CAATATCTCGATACATACATACATACATACATACATACATACATACATACATACAT 381
QY 681 GCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 740
Db 380 ACTAGCAACCGACAACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 322
QY 741 ACCCTTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 800
Db 321 ACCCTTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 262
QY 801 CGCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 845
Db 261 CGCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 217

```

Search completed: December 28, 2002, 02:59:52
Job time : 185.866 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 28, 2002, 02:47:28 ; Search time 1119.58 Seconds
(without alignments)
12237.926 Million cell updates/sec

Title: US-09-709-103-1

Perfect score: 846

Sequence: 1 atgaactgcgcgcatgat.....agcgtgcgcacgtag 846

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : EST:*

- 1: em_estba:*
- 2: em_esthum:*
- 3: em_estlin:*
- 4: em_estmu:*
- 5: em_estov:*
- 6: em_estpl:*
- 7: em_estro:*
- 8: em_hic:*
- 9: gb_est1:*
- 10: gb_est2:*
- 11: gb_hic:*
- 12: gb_est3:*
- 13: gb_est4:*
- 14: gb_est5:*
- 15: em_estfun:*
- 16: em_estom:*
- 17: gb_gss:*
- 18: em_gss_hum:*
- 19: em_gss_inv:*
- 20: em_gss_pln:*
- 21: em_gss_vrt:*
- 22: em_gss_fun:*
- 23: em_gss_mam:*
- 24: em_gss_mus:*
- 25: em_gss_othr:*
- 26: em_gss_pro:*
- 27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	695.8	82.2	1035	14	BM919341
2	669.8	79.2	962	13	BM543472
3	664.8	78.6	965	9	AL533318
4	661.8	78.2	742	10	AW028127
5	650.4	76.9	958	14	BQ719566
6	641	75.8	1103	14	BM920514

7	612.6	72.4	1053	14	BM921737
8	586	69.3	1023	13	BM543630
9	580	68.6	1032	14	BQ067637
10	540.8	63.9	758	13	BI553776
11	536.2	63.4	814	13	BI596688
12	521.4	61.6	1137	14	BM921656
13	521.4	61.6	1160	14	BM805574
14	519.6	61.4	831	13	BI596637
15	513.6	60.7	655	12	BG711792
16	513	60.6	649	12	BG085090
17	506	59.8	506	13	BM311047
18	481	56.9	699	12	BG706012
19	470	55.6	700	13	BI596509
20	469	55.4	699	12	BG709229
21	448	53.0	674	13	BI601563
22	446	52.7	617	13	BI393669
23	439	51.9	726	12	BF613135
24	438.8	51.9	675	13	BI545172
25	433.4	51.2	666	13	BI549939
26	427.8	50.6	644	13	BG969048
27	426.4	50.4	689	10	BG632699
28	417.6	49.4	589	13	BM426066
29	415.8	49.1	648	10	BG636889
30	401.6	47.5	571	13	BI682922
31	401	47.4	1006	14	BQ073742
32	399	47.2	947	14	BQ954076
33	389.6	46.1	904	14	BQ947936
34	386.4	45.7	1438	14	BQ807669
35	384.2	45.4	888	13	BI754083
36	374.2	44.2	640	13	BJ490945
37	365.6	43.2	505	13	BI391213
38	362.8	42.9	509	10	AW915326
39	359	42.4	1332	13	BM460899
40	358	42.3	1300	11	AK015898
41	354	41.8	456	12	BF555582
42	351	41.5	556	12	BG654362
43	349.6	41.3	710	13	BJ526038
44	342	40.4	1404	13	BM548645
45	341	40.3	531	13	BI823596

ALIGNMENTS

RESULT 1
BM919341
LOCUS
DEFINITION
AGENCOURT_6715681 NIH_MGC_120 Homo sapiens cDNA clone IMAGE:5748585
5', mRNA sequence.
BM919341
ACCESSION
BM919341
VERSION
BM919341.1 GI:19369720
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens

1035 bp mRNA linear EST 12-MAR-2002

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

NIH-MGC http://mgc.nci.nih.gov/

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabps-femail.nih.gov

Tissue Procurement: Life Technologies, Inc.

cDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM12776 row: j column: 10

High quality sequence stop: 658.

Location/Qualifiers

1. .1035

157

Db	524	CTGGTGTTCAGTCTGGACAAACCGGACTCTTCGAGGAGGTGTCAGCGCTCAGGACGAG	583
QY	361	ATCCTCGACACCAAGTCTTGGCTCAAGAACAAACCAAGGAGAACGTTGGAGCTTGCCTCG	420
Db	584	ATCCTCGACACCAAGTCTTGGCTCAAGAACAAACCAAGGAGAACGTTGGAGCTTGCCTCG	643
QY	421	GTATCTGGGGCAACCAAGGTTGACCGGACTTCTACCGGAGGTGGACCAAGCGAGATC	480
Db	644	GTATCTGGGGCAACCAAGGTTGACCGGACTTCTACCGGAGGTGGACCAAGCGAGATC	703
QY	481	GAGCAGCTGGGGGAGACACCGGCTTGGCTTCTACCGGAGGTGGACCAAGCGAGATC	540
Db	704	GAGCAGCTGGGGGAGACACCGGCTTGGCTTCTACCGGAGGTGGACCAAGCGAGATC	763
QY	541	AACAGCAGCTGGGACGAGTCTTCCGGCGCTTTCGCCATGTCGACCAAGTGGCCCGGAG	600
Db	764	AACAGCAGCTGGGACGAGTCTTCCGGCGCTTTCGCCATGTCGACCAAGTGGCCCGGAG	823
QY	601	ATGAGCCGACGCTGACACCAAGTCTTGGCTTCTACCGGAGGTGGACCAAGCGAGATC	660
Db	824	ATGAGCCGACGCTGACACCAAGTCTTGGCTTCTACCGGAGGTGGACCAAGCGAGATC	883
QY	661	CGGCTGC-GGAACAGAGAGTCTTGGGGCGG 692	
Db	884	CGGCTGC-GGAACAGAGAGTCTTGGGGCGG 916	
RESULT 3			
AL533318			
LOCUS	AL533318	LTI_FL015_Brn1	965 bp mRNA linear EST 13-FEB-2001
DEFINITION			Prime, mRNA sequence.
ACCESSION	AL533318		
VERSION	AL533318.1	GI:12796811	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.		
TITLE	Li, W.B., Gruber, C., Jesse, J., and Polayes, D.		
JOURNAL	Full-length cDNA libraries and normalization		
COMMENT	Unpublished (2001)		
FEATURES	Genoscope - Centre National de Sequencage		
source	Bp 191 91006 EVRI cedex - France		
	Email: segre@genoscope.cns.fr, Web: www.genoscope.cns.fr.		
	Location/Qualifiers		
	1..965		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="CS0DN003YJ19"		
	/clone_lib="LTI_FL015_Brn1"		
	/sex="male"		
	/tissue_type="Adult brain"		
	/note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA		
	was primed with a NotI-oligo(dT) primer. Five prime end		
	enriched, double-stranded cDNA was digested with Not I and		
	cloned into the Not I and Eco RV sites of the pCMVSPORT 6		
	vector. Library was constructed by Life Technologies		
	Contact: Feng Liang Life Technologies, a division of		
	Invitrogen 9800 Medical Center Drive Rockville, Maryland		
	20850, USA Fax: (1) 301 610 8371 Email:		
	fliang@lifetech.com URL:		
	http://fulllength.invitrogen.com"		
BASE COUNT	189 a 352 c 273 g 144 t	7 others	
ORIGIN			
Query Match	78.6%; Score 664.8; DB 9; Length 965;		
Best Local Similarity	95.5%; Pred. No. 1.2e-120;		
Matches 720; Conservative	7; Mismatches 19; Indels 8; Gaps 4;		

QY	1	ATGAAGTGGCGGATGATCAAGAAGATGTCGCCGAGGAGCTCGAGGCTGAGTATCCCG	60
Db	191	ATGAAGTGGCGGATGATCAAGAAGATGTCGCCGAGGAGCTCGAGGCTGAGTATCCCG	250
QY	61	GCCAAAGTGGCTATCGCATGTCATCCCTCGCTCGCTCAAGGTGGCAAGCGCCATC	120
Db	251	GCCAAAGTGGCTATCGCATGTCATCCCTCGCTCGCTCAAGGTGGCAAGCGCCATC	310
QY	121	GTGTCCGGCTTCTCACCGCGGCTTCGAGGAGGCTTACAGCCTTACATCGAGGACTTC	180
Db	311	GTGTCCGGCTTCTCACCGCGGCTTCGAGGAGGCTTACAGCCTTACATCGAGGACTTC	370
QY	181	CACCGCAAGTTCATCCATCCGCGGAGGCTTACAGGCTCGACATCTTCGACAGTCC	240
Db	371	CACCGCAAGTTCATCCATCCGCGGAGGCTTACAGGCTCGACATCTTCGACAGTCC	300
QY	241	GGCAACACCCCTTCCCGGCTTCGCGGCTTCCTCCATCTTCACAGGAGGCTTTTCATC	300
Db	431	GGCAACACCCCTTCCCGGCTTCGCGGCTTCCTCCATCTTCACAGGAGGCTTTTCATC	489
QY	301	CTGGTGTTCAGTCTGGCAACCGGCTTCCTTCGAGGAGGCTTCAGGCGGCTTCAGGAGC	360
Db	490	CTGGTGTTCAGTCTGGCAACCGGCTTCCTTCGAGGAGGCTTCAGGCGGCTTCAGGAGC	549
QY	361	ATCCTCGACACCAAGTCTTCGCTCAAGAACAAACCAAGGAGGAGGCTTCGAGATCT	420
Db	550	ATCCTCGACACCAAGTCTTCGCTCAAGAACAAACCAAGGAGGAGGCTTCGAGATCT	509
QY	421	GTATCTCGGCAACAAAGGCTTCGCGGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCG	600
Db	610	GTATCTCGGCAACAAAGGCTTCGCGGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCG	669
QY	481	GAGCAGTGTGGGCGGAGGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCT	540
Db	670	GAGCAGTGTGGGCGGAGGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCT	729
QY	541	AACAGCAGCTTCGACAGATGTTCCGCGGCTTCGCTTCGCTTCGCTTCGCTTCGCT	600
Db	730	AACAGCAGCTTCGACAGATGTTCCGCGGCTTCGCTTCGCTTCGCTTCGCTTCGCT	789
QY	601	ATGAGCCGACCTTCGACAGGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCG	660
Db	790	ATGAGCCGACCTTCGACAGGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCG	848
QY	661	CGCTCGGCAACAAAGGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCG	720
Db	849	CGCTCGGCAACAAAGGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCG	902
QY	721	GAGCGCTTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCG	754
Db	903	GAGCGCTTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCG	936
RESULT 4			
AL533318			
LOCUS	AL533318	LTI_FL015_Brn1	742 bp mRNA linear EST 27-OCT-1999
DEFINITION			similar to TR:O35626 O35626 RAS, DEXAMETHASONE-INDUCED 1 ;, MRNA
ACCESSION	AL533318		
VERSION	AL533318.1	GI:5886883	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.		
TITLE	NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.		
JOURNAL	National Cancer Institute, Cancer Genome Anatomy Project (CGAP),		
COMMENT	Unpublished (1997)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgaabp@mail.nih.gov		

Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution Information can be found through the I.M.A.G.E. Consortium/LLNL at:
 www.blo.llnl.gov/bbrp/image/image.html
 Seq primer: -40UP from Gibco
 High quality sequence stop: 445.
 Location/Qualifiers

FEATURES
 source

1. 742
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2530668"
 /clone_lib="NCI_CGAP_Kid11"
 /lab_host="DH10B"
 /note="Organ: kidney; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; plasmid DNA from the normalized library NCI_CGAP_Kid3 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneids 132376-132391, 1456007-1456775, and 1500552-1502855). Subtraction by Bento Soares and M. Fatima Bonaldo."
 BASE COUNT 149 a 260 c 208 g 122 t 3 others

Query Match 78.2%; Score 661.8; DB 10; Length 742;
 Best Local Similarity 97.3%; Pred. No. 4.4e-120;
 Matches 672; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 1 ATGAAGCTGGCGGATGATCAAGAGATGTGCCGAGCGACTCGGAGTGTATCCCG 60
 Db 35 ATGAAGCTGGCGGATGATCAAGAGATGTGCCGAGCGACTCGGAGTGTATCCCG 94
 QY 61 GCCAAGCTGCTATCCATCCCGCGAGGTCTTCCAAAGTGGCGAAGACGGCCATC 120
 Db 95 GCCAAGCTGCTATCCATCCCGCGAGGTCTTCCAAAGTGGCGAAGACGGCCATC 154
 QY 121 GTGTGGCGCTTCTTCCATCCCGCGAGGTCTTCCAAAGTGGCGAAGACGGCCATC 180
 Db 155 GTGTGGCGCTTCTTCCATCCCGCGAGGTCTTCCAAAGTGGCGAAGACGGCCATC 214
 QY 181 CACCGCAAGTCTTCTTCCATCCCGCGAGGTCTTCCAAAGTGGCGAAGACGGCCATC 240
 Db 215 CACCGCAAGTCTTCTTCCATCCCGCGAGGTCTTCCAAAGTGGCGAAGACGGCCATC 274
 QY 241 GGCACACACCGCTTCTTCCATCCCGCGAGGTCTTCCAAAGTGGCGAAGACGGCCATC 300
 Db 275 GGCACACACCGCTTCTTCCATCCCGCGAGGTCTTCCAAAGTGGCGAAGACGGCCATC 334
 QY 301 CTGGTGTTCAGTCTGGACACCGGACTCTTCCGAGAGGTGTCAGCGGCTCAGGACGACG 360
 Db 335 CTGGTGTTCAGTCTGGACACCGGACTCTTCCGAGAGGTGTCAGCGGCTCAGGACGACG 394
 QY 361 ATCTCGACACCAAGTCTTCCGCTCAAGAACCAAAACCAAGGAGACGTGGAGCGCCCTG 420
 Db 395 ATCTCGACACCAAGTCTTCCGCTCAAGAACCAAAACCAAGGAGACGTGGAGCGCCCTG 454
 QY 421 GTCATCTGGGGCAACAGGGTACCGGACTCTTCCGAGAGGTGACCGAGCGAGATC 480
 Db 455 GTCATCTGGGGCAACAGGGTACCGGACTCTTCCGAGAGGTGACCGAGCGAGATC 514
 QY 481 GACAGCTGTGGGACGACGACCGGCTTCCGAGAGGTGACCGAGCGAGATC 540
 Db 515 GACAGCTGTGGGACGACGACCGGCTTCCGAGAGGTGACCGAGCGAGATC 574
 QY 541 AACAGCAGCTGGACGAGATGTTCCGCGGCTTTCGCGATGCGCAAGCTGCCAGCGAG 600
 Db 575 AACAGCAGCTGGACGAGATGTTCCGCGGCTTTCGCGATGCGCAAGCTGCCAGCGAG 634

QY 601 ATGAGCCAGACCTGCACCGAAGTCTCGGTGCACTACTGCGAGCTGTGCACAGAAG 660
 Db 635 ATGAGCCAGACCTGCACCGAAGTCTCGGTGCACTACTGCGAGCTGTGCACAGAAG 694
 QY 661 GCGCTGCGGAACAGAAAGCTGCTGCGGCGG 691
 Db 695 GCGCTGCGGAACAGAAAGCTGCTGCGGCTCGG 725

RESULT 5

BQ719566

LOCUS

DEFINITION

AGENCOURT_8305191 Lupski_sympathetic_trunk Homo sapiens cDNA clone

IMAGE:6193170 5', mRNA sequence.

ACCESSION

BQ719566

VERSION

BQ719566.1

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE

1 (bases 1 to 958)

AUTHORS

NIH-MGC http://mgi.nci.nih.gov/.

TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgabbs-remail.nih.gov

Tissue Procurement: Dr. James R. Lupski

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM13596 row: f column: 19

High quality sequence stop: 514.

Location/Qualifiers

1. 958

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:6193170"

/clone_lib="Lupski_sympathetic_trunk"

/sex="male"

/tissue_type="sympathetic trunk"

/dev_stage="adult, 16 yr"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6 (Life Technologies); Site_1:

NotI; Site_2: SalI; cDNA made by oligo-dT priming.

Directionally cloned using the following adaptors:

5'-TCGACCCACGCTCCG-3' and

5'-GACTAGTTCTAGATCGGAGCGGCGCCCT(15)-3'. Size selected >

1 kb for average insert length 1.9 kb. This is a primary

library, non-amplified. Library constructed by Life

Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor

College of Medicine); available through Life

Technologies."

BASE COUNT 185 a 339 c 275 g 159 t

ORIGIN

Query Match

Best Local Similarity

Matches 684; Conservative

76.9%; Score 650.4; DB 14; Length 958;

95.0%; Pred. No. 7.8e-118;

0; Mismatches 31; Indels 5; Gaps 1;

QY 43 TCGGAGCTGAGTATCCGCGCCCAAGAACTGCTATCGCATGGTTCATCTCGGCTCGTCCAAG 102

Db 1 TCGGAGCTGAGTATCCGCGCCCAAGAACTGCTATCGCATGGTTCATCTCGGCTCGTCCAAG 60

QY 103 GTGGGCAAGACGGCCATCGTGTGCGGCTTCTTCCACCGCGCGCTTCGAGGACGCTACACG 162

Db 61 GTGGGCAAGACGGCCATCGTGTGCGGCTTCTTCCACCGCGCGCTTCGAGGACGCTACACG 120

QY 163 CCTACCATCGAGGACTTCCACCGCAGTTCTACTCCATCCGCGCGAGTCTTACCAGCTC 222

```

Db 121 CCTACATCGAGGACTTCCACCGCAAGTTCTATCTCCATCCGCGGAGGTCTACAGCTC 180
QY 223 GACATCTCGACACGTCGCGCAACACCGTTCCCGCGCATGCGCGGCTCTCATCTC 282
Db 181 GACATCTCGACACGTCGCGCAACACCGTTCCCGCGCATGCGCGGCTCTCATCTC 240
QY 283 ACAGGAGAGGTTTTCATCTCGTGTGTAGTCTGGACAAACCGGACTCTCTCGAGGAGTG 342
Db 241 ACAGGAGAGGTTTTCATCTCGTGTGTAGTCTGGACAAACCGGACTCTCTCGAGGAGTG 300
QY 343 CAGCGGCTCAGGACGAGATCTCGACACCAAGTCTTGCCCTCAGAACACCGGAG 402
Db 301 CAGCGGCTCAGGACGAGATCTCGACACCAAGTCTTGCCCTCAGAACACCGGAG 360
QY 403 AACGTGGAGTCCCGCTCGTATCTGGCGCAACAGGAGTACCGGAGTCTTACCGGAG 462
Db 361 AACGTGGAGTCCCGCTCGTATCTGGCGCAACAGGAGTACCGGAGTCTTACCGGAG 420
QY 463 GTGGACACGCGGAGATCGAGCAGCTGTGGGCGACACCGGAGTACCGGAGTCTTACCGGAG 522
Db 421 GTGGACACGCGGAGATCGAGCAGCTGTGGGCGACACCGGAGTACCGGAGTCTTACCGGAG 480
QY 523 GAGATCTCGGCAAGACAGCAGCTGTGGGCGACACCGGAGTACCGGAGTCTTACCGGAG 582
Db 481 GAGATCTCGGCAAGACAGCAGCTGTGGGCGACACCGGAGTACCGGAGTCTTACCGGAG 540
QY 583 GCCAAGTCCCGGAGATCGAGCAGCTGTGGGCGACACCGGAGTACCGGAGTCTTACCGGAG 642
Db 541 GCCAAGTCCCGGAGATCGAGCAGCTGTGGGCGACACCGGAGTACCGGAGTCTTACCGGAG 600
QY 643 GAGTGTGTGCAAGAGGCGTGTGGGCGACACCGGAGTACCGGAGTCTTACCGGAG 697
Db 601 GAGTGTGTGCAAGAGGCGTGTGGGCGACACCGGAGTACCGGAGTCTTACCGGAG 660
QY 698 GCGGCGGCGGCGGAGTGTGGGCGACACCGGAGTACCGGAGTCTTACCGGAG 757
Db 661 GCGGCGGCGGCGGAGTGTGGGCGACACCGGAGTACCGGAGTCTTACCGGAG 720

```

```

RESULT 6
BM920514
LOCUS BM920514 1103 bp mRNA linear EST 12-MAR-2002
DEFINITION AGENCOURT_6709473 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5750507
5', mRNA sequence.
ACCESSION BM920514
VERSION BM920514.1 GI:19370893
KEYWORDS EST.
SOURCE Human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing Arrayed by: The I.M.A.G.E. Consortium (LLNL)
Clone Distribution: Agencourt Bioscience Corporation
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12781 row: j column: 12
High quality sequence stop: 676.
Location/Qualifiers
1..1103
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5750507"
/clone_lib="NIH_MGC_122"
FEATURES
source

```

```

/lab_host="DH10B"
/note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6;
Site_1: NotI; Site_2: EcoRV (destroyed); RNA source
anonymous pool of 24 week female lung, 16 week female
spleen, and 20-22 week male spleens. Library is oligo-dT
primed and directionally cloned (EcoRV site is oligo-dT
upon cloning). Average insert size 1.4 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(invitrogen). Research Genetics tracking code 026. Note:
this is a NIH_MGC Library."
BASE COUNT 230 a 402 c 314 g 156 t
ORIGIN 1 others

```

```

Query Match 75.8%; Score 641; DB 14; Length 1103;
Best Local Similarity 90.7%; Pred. No. 5.5e-116;
Matches 764; Conservative 0; Mismatches 61; Indels 17; Gaps 7;
QY 1 ATGAACCTGGCGGATGATCAAGAGATGTGCCGAGCGACTCGGAGCTGAGTATCCCG 60
Db 201 ATGAACCTGGCGGATGATCAAGAGATGTGCCGAGCGACTCGGAGCTGAGTATCCCG 260
QY 61 GCCAAGAACTGCTATCGCATGGTTCATCTCGGCTCTCAAGGTGGGCAAGCGGCATC 120
Db 261 GCCAAGAACTGCTATCGCATGGTTCATCTCGGCTCTCAAGGTGGGCAAGCGGCATC 320
QY 121 GTGTGCGGCTTCTACCGGCGCTTTCGAGGACGCTTACACGCTTACCATCGGAGCTTC 180
Db 321 GTGTGCGGCTTCTACCGGCGCTTTCGAGGACGCTTACACGCTTACCATCGGAGCTTC 380
QY 181 CACGCAAGTCTTACTTCATCCCGCGGAGGTCTACAGCTCGACATCTCGACACGTC 240
Db 381 CACGCAAGTCTTACTTCATCCCGCGGAGGTCTACAGCTCGACATCTCGACACGTC 440
QY 241 GGCACACACCGTTCCTCCCGCATTCGCGGCGCTTCCATCTCACAGGAGAGCTTTTCATC 300
Db 441 GGCACACACCGTTCCTCCCGCATTCGCGGCGCTTCCATCTCACAGGAGAGCTTTTCATC 500
QY 301 CTGCTGTCAGTCTGGACACCGGAGTCTTCGAGGAGGTGCGGCTCAGGCGGAGATC 360
Db 501 CTGCTGTCAGTCTGGACACCGGAGTCTTCGAGGAGGTGCGGCTCAGGCGGAGATC 560
QY 361 ATCTCTGACACCAAGTCTTTCCTCAAGAACAAACAAAGGAGAGCTGCGGCTG 420
Db 561 ATCTCTGACACCAAGTCTTTCCTCAAGAACAAACAAAGGAGAGCTGCGGCTG 620
QY 421 GTCTATCTGGGCAACAGGAGTTCACCGGAGTCTACCGGAGGTGACCGGAGATC 480
Db 621 GTCTATCTGGGCAACAGGAGTTCACCGGAGTCTACCGGAGGTGACCGGAGATC 680
QY 481 GAGCAGTGTGGGCGGAGCGACCGGCTTTCGAGATCTTCGAGATCTGGGCGGAGAG 540
Db 681 GAGCAGTGTGGGCGGAGCGACCGGCTTTCGAGATCTTCGAGATCTGGGCGGAGAG 740
QY 541 AACAGCAGCTGGGAGAGATGTCGCGGCTTTCGAGATCTTCGAGATCTGGGCGGAGAG 599
Db 741 AA-AGCAGCTGTGGGCGGAGTTCGCGGCTTTCGAGATCTTCGAGATCTGGGCGGAGAG 799
QY 600 GATGAGCGGAGAGCTTCACCGGAGGTCTCGGTCAGTCTGCGAGTCTGCGACAGAA 659
Db 800 GATGAGCGGAGAGCTTCACCGGAGGTCTCGGTCAGTCTGCGAGTCTGCGACAGAA 859
QY 660 GCGGC-TGCCGAGCAACAA-----GCTGCTCGGCGGCGGAGCGGCGGCGGCGGC 711
Db 860 AGGCTTTTTCGAGAGAGAGAGTGTTCGCGGCGGCGGAGCGGCGGCGGCGGCGGC 919
QY 712 GACCGCGGCGGAGCGCTTTTGGCATCTGCGGAGCGCTTTCGCGGCGGCGGCGGCGGC 968
Db 920 ACCGCGGCGGAGAGCGCTTTTGGCATCTGCGGAGCGCTTTCGCGGCGGCGGCGGCGGC 979
QY 769 AGCG-ACCTCATGTACATCGGCGGAGAGCGC-----AGCGCGGCGGAGCGGCGGCGGCGGC 824
Db 980 AGGGAACCTTATGGAAATCCCGCAAGAGAGCGCGGCGGCGGCGGCGGCGGCGGC 1039

```

QY 241 GGCAACACCCGTTCCCGCCATGCGCGCCTCTCATCTCACAGGAGACGTTTTCATC 300
 Db 440 GGCAACACCCGTTCCCGCCATGCGCGCCTCTCATCTCACAGGAGACGTTTTCATC 499
 QY 301 CTGGTGTTCAGTCTGGACAAACCCGAGCTCTTCGAGGAGGTGACGGCTCAGGAGCAG 360
 Db 500 CTGGTGTTCAGTCTGGACAAACCCGAGCTCTTCGAGGAGGTGACGGCTCAGGAGCAG 559
 QY 361 ATCTCTGACACCAAGTCTTCCCTCAAGAAACAAACCAAGAGAACTGGACGTGCCCTG 420
 Db 560 ATCTCTGACACCAAGTCTTCCCTCAAGAAACAAACCAAGAGAACTGGACGTGCCCTG 619
 QY 421 GTCTATCTGCGGCAACAAAGGTGACCGGAGCTTCTACCGGAGGTGGACCGGAGATC 480
 Db 620 GTCTATCTGCGGCAACAAAGGTGACCGGAGCTTCTACCGGAGGTGGACCGGAGATC 679
 QY 481 GAGCAGCTGTGGGCGAGACGCCCGGAGCTTCTGCGGAGGTTCGAGATCTCGGCAAGAG 540
 Db 680 GAGCAGCTGTGGGCGAGACGCCCGGAGCTTCTGCGGAGGTTCGAGATCTCGGCAAGAG 739
 QY 541 AACAGCAG-CTTGGACACAGATGTTCCCGGCGCTCTTTCGCCATGCCCAAGCTGCCAGCGA 599
 Db 740 AACAGCAGCTTGGACACAGATGTTCCCGGCGCTCTTTCGCCATGCCCAAGCTGCCAGCGA 799
 QY 600 GATGAGCCGAGCTTGGACACAGATGTTCCCGGCGCTCTTTCGCCATGCCCAAGCTGCCAGCGA 859
 Db 800 GATGAGCCGAGCTTGGACACAGATGTTCCCGGCGCTCTTTCGCCATGCCCAAGCTGCCAGCGA 919
 QY 657 GAAGGCGCTTGGACACAGATGTTCCCGGCGCTCTTTCGCCATGCCCAAGCTGCCAGCGA 716
 Db 860 AAAAGGGCTGCGGGAACAAAAAACTGCTGCGGGGCGGTAACCCGTTGGCGGCGCG 919
 QY 717 GCGCGAC 723
 Db 920 GAGGAAC 926

RESULT 8
 BM543630
 LOCUS DEFINITION 1023 bp mRNA linear EST 20-FEB-2002
 AGENCOURT_6492527 NIH_MGC_124 Homo sapiens cDNA clone IMAGE:5726764
 5', mRNA sequence.
 BM543630
 ACCESSION
 VERSION BM543630.1 GI:18774186
 KEYWORDS
 SOURCE human.
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 (bases 1 to 1023)
 NIH-MGC http://mgs.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT: Robert Strausberg, Ph.D.
 Email: coapbs-remail.nih.gov
 Tissue Procurement: Invitrogen
 CDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L1AM12719 row: m column: 05
 High quality sequence stop: 637.
 Location/Qualifiers
 1..1023
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="IMAGE:5726764"
 /clone_lib="NIH_MGC_124"
 /issue_type="hippocampus"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pcMV-SpO16; site_1: EcoRV"

FEATURES
 source
 1..1023
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="IMAGE:5726764"
 /clone_lib="NIH_MGC_124"
 /issue_type="hippocampus"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pcMV-SpO16; site_1: EcoRV"

QY 825 GG 826
 Db 1040 AG 1041
 RESULT 7
 BM921737 1053 bp mRNA linear EST 12-MAR-2002
 LOCUS DEFINITION AGENCOURT_6708101 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5753301
 5', mRNA sequence.
 BM921737
 ACCESSION
 VERSION BM921737.1 GI:19372116
 KEYWORDS
 SOURCE human.
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 (bases 1 to 1053)
 NIH-MGC http://mgs.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT: Robert Strausberg, Ph.D.
 Email: coapbs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 CDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L1AM12788 row: n column: 22
 High quality sequence stop: 600.
 Location/Qualifiers
 1..1053
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="IMAGE:5753301"
 /clone_lib="NIH_MGC_115"
 /lab_host="DH10B"
 /note="Organ: pooled brain, lung, testis; Vector: pcMV-SpO16; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27; 1 male lung, age 27; and 1 male testis, age 69. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.8 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 021. Note: this is a NIH_MGC Library."

BASE COUNT 215 a 390 c 297 g 150 t
 ORIGIN
 Query Match 72.48; Score 612.6; DB 14; Length 1053;
 Best Local Similarity 92.08; Pred. No. 2.le-110;
 Matches 669; Conservative 0; Mismatches 54; Indels 4; Gaps 2;
 QY 1 ATGAAGTGGCGCGATGATCAAGAAGATGCGCGAGCGACTCGGAGCTGAGTATCCCG 60
 Db 200 ATGAAGTGGCGCGATGATCAAGAAGATGCGCGAGCGACTCGGAGCTGAGTATCCCG 259
 QY 61 GCGAAGTGTATGCGATGTCATCTCGGCTGCTCAAGGTGGGCAAGACGGCCATC 120
 Db 260 GCGAAGTGTATGCGATGTCATCTCGGCTGCTCAAGGTGGGCAAGACGGCCATC 319
 QY 121 GTGTGCGGTCTCTCACCGCGCGCTTCGAGGACGCGCTACAGCGCTACGAGGATTC 180
 Db 320 GTGTGCGGTCTCTCACCGCGCGCTTCGAGGACGCGCTACAGCGCTACGAGGATTC 379
 QY 181 CACCGAAGTTTACTTCGATCGCGCGAGGTCTACAGCTCGACATCCTCGACACGTC 240
 Db 380 CACCGAAGTTTACTTCGATCGCGCGAGGTCTACAGCTCGACATCCTCGACACGTC 439

(destroyed); Site_2: NotI; RNA source male hippocampus, age 27. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 0.9-4 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 012.*

BASE COUNT 196 a 378 c 281 g 168 t
ORIGIN

Query Match 69.3%; Score 586; DB 13; Length 1023;
Best Local Similarity 97.9%; Pred. No. 3.4e-105;
Matches 615; Conservative 0; Mismatches 10; Indels 3; Gaps 2;

```

QY 1 ATGAAGTGGCGGATCATCAAGAGATGTGCCGAGGACGCTCGAGCTGAGTATCCCG 60
   |||||
Db 149 ATGAAGTGGCGGATCATCAAGAGATGTGCCGAGGACGCTCGAGCTGAGTATCCCG 208
QY 61 GCCAAGAACTGCTATCGCATGCTCATCTCGGCTCGTCCAAAGTGGGCAAGAGGCCATC 120
   |||||
Db 209 GCCAAGAACTGCTATCGCATGCTCATCTCGGCTCGTCCAAAGTGGGCAAGAGGCCATC 268
QY 121 GTGTGGCGCTTCCTCACCGGCGGCTTCGAGGAGCGCTACACGCCCTACCATCGAGGACTTC 180
   |||||
Db 269 GTGTGGCGCTTCCTCACCGGCGGCTTCGAGGAGCGCTACACGCCCTACCATCGAGGACTTC 328
QY 181 CACCGCAAGTTTACTTCCATCCCGCGGAGGTTTACCAGCTCGACATCTCGACACGTC 240
   |||||
Db 329 CACCGCAAGTTTACTTCCATCCCGCGGAGGTTTACCAGCTCGACATCTCGACACGTC 388
QY 241 GGCACACACCGTTTCCCGCGGCGGCTTCGAGGAGCGCTTCATCTCCACAGGAGAGCTTTTCATC 300
   |||||
Db 389 GGCACACACCGTTTCCCGCGGCGGCTTCGAGGAGCGCTTCATCTCCACAGGAGAGCTTTTCATC 448
QY 301 CTGTGTTCAGTCTGACAAACCGGAGCTTCCTCGAGGAGGTGCGAGGCTCGAGCGAG 360
   |||||
Db 449 CTGTGTTCAGTCTGACAAACCGGAGCTTCCTCGAGGAGGTGCGAGGCTCGAGCGAG 508
QY 361 ATCTCGACACCAAGTCTTGCCTCAAGAACAAACAAAGGAGAACTGGAGCTGCCCTG 420
   |||||
Db 509 ATCTCGACACCAAGTCTTGCCTCAAGAACAAACAAAGGAGAACTGGAGCTGCCCTG 568
QY 421 GTCATCTGGGCAACAAAGGTTGACCGGAGCTTCTACCGGAGGTGGACGCGGAGATC 480
   |||||
Db 569 GTCATCTGGGCAACAAAGGTTGACCGGAGCTTCTACCGGAGGTGGACGCGGAGATC 628
QY 481 GAGCAGCTGGTGGGCGAGACCGCGGCTGCGCTTCTTCCAGATCTCGGCAAGAG 540
   |||||
Db 629 GAGCAGCTGGTGGGCGAGACCGCGGCTGCGCTTCTTCCAGATCTCGGCAAGAG 688
QY 541 AACAGCAGCT--GACACAGATCTTCCGCGGCTTCTTCCAGATCTCGGCAAGAG 597
   |||||
Db 689 AACAGCAGCTGGGACAGATGTTCCGCGGCTTCTTCCAGATCTCGGCAAGAG 748
QY 598 GAGATGAGCCAGACCTGACCGGCAAG 625
   |||||
Db 749 GAGATGAGCCAGACCTGACCGGCAAG 776

```

RESULT 9
BQ067637
LOCUS BQ067637 1032 bp mRNA linear EST 02-APR-2002
DEFINITION AGENCOURT_6759053 NIH_MGC_115 Homo sapiens cdna clone IMAGE:5755214
5', mRNA sequence.

ACCESSION BQ067637
VERSION BQ067637.1 GI:19896683
KEYWORDS EST.

SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1032)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE JOURNAL COMMENT

National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapb-re@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cdNA Library Preparation: The I.M.A.G.E. Consortium (LLNL)
cdNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM12793 row: n column: 15
High quality sequence stop: 602.

Location/Qualifiers
1. .1032
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5755214"
/clone_lib="NIH_MGC_115"
/lab_host="DH10B"

FEATURES source

/note="Organ: pooled brain, lung, testis; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27; 1 male lung, age 27; and 1 male testis, age 69. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.8 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed and Gruber (Invitrogen). Research Genetics tracking code 021. Note: this is a NIH_MGC Library."

BASE COUNT 210 a 375 c 298 g 148 t 1 others
ORIGIN

Query Match 66.6%; Score 580; DB 14; Length 1032;
Best Local Similarity 91.6%; Pred. No. 5.1e-104;
Matches 683; Conservative 0; Mismatches 50; Indels 13; Gaps 6;

```

QY 1 ATGAAGTGGCGGATCATCAAGAGATGTGCCGAGGACGCTCGAGCTGAGTATCCCG 60
   |||||
Db 201 ATGAAGTGGCGGATCATCAAGAGATGTGCCGAGGACGCTCGAGCTGAGTATCCCG 260
QY 61 GCCAAGAACTGCTATCGCATGCTCATCTCGGCTCGTCCAAAGTGGGCAAGAGGCCATC 120
   |||||
Db 261 GCCAAGAACTGCTATCGCATGCTCATCTCGGCTCGTCCAAAGTGGGCAAGAGGCCATC 320
QY 121 GTGTGGCGCTTCCTCACCGGCGGCTTCGAGGAGCGCTTACAGCTTACCATCGAGGACTTC 180
   |||||
Db 321 GTGTGGCGCTTCCTCACCGGCGGCTTCGAGGAGCGCTTACAGCTTACCATCGAGGACTTC 380
QY 181 CACCGCAAGTTTACTTCCATCCCGCGGAGGTTTACCAGCTCGACATCTCGACACGTC 240
   |||||
Db 381 CACCGCAAGTTTACTTCCATCCCGCGGAGGTTTACCAGCTCGACATCTCGACACGTC 440
QY 241 GGCACACACCGTTTCCCGCGGCTTCGAGGAGGTGCGAGGCTCGAGCGGCTCGAGCGAG 500
   |||||
Db 441 GGCACACACCGTTTCCCGCGGCTTCGAGGAGGTGCGAGGCTCGAGCGGCTCGAGCGAG 560
QY 301 CTGTGTTCAGTCTGACAAACCGGAGCTTCCTCGAGGAGGTGCGAGGCTCGAGCGAG 360
   |||||
Db 501 CTGTGTTCAGTCTGACAAACCGGAGCTTCCTCGAGGAGGTGCGAGGCTCGAGCGAG 560
QY 361 ATCTCTGACACCAAGTCTTGCCTCAAGAACAAACAAAGGAGAACTGGAGCTGCCCTG 420
   |||||
Db 561 ATCTCTGACACCAAGTCTTGCCTCAAGAACAAACAAAGGAGAACTGGAGCTGCCCTG 620
QY 421 GTCATCTGGGCAACAAAGGTTGACCGGAGCTTCTTACCGGAGGTGGACGCGGAGATC 480
   |||||
Db 621 GTCATCTGGGCAACAAAGGTTGACCGGAGCTTCTTACCGGAGGTGGACGCGGAGATC 680
QY 481 GAGCAGCTGGTGGGCGAGACCGCGGCTGCGCTTCTTCCAGATCTCGGCAAGAG 540
   |||||
Db 681 GAGCAGCTGGTGGGCGAGACCGCGGCTGCGCTTCTTCCAGATCTCGGCAAGAG 740

```


Mon Dec 30 09:16:52 2002

352 AGGAGCAGATCTCGACACCAAGTCTTGCTCAAGAACAAACAAACCA-GGAGACGTGA 410
 66 AGGAGCAGATCTCGACACCAAGTCTTGCTCAAGAACAAACAAACCAAGGAGAGTGA 125
 411 CGTGCCTCTGTCATCTGCGGCAACAAAGGTGACCGGACTTCTACCGGAGGTGACCA 470
 126 CGTGCCTCTGTCATCTGCGGCAACAAAGGTGACCGGACTTCTACCGGAGGTGACCA 185
 471 GCGGAGATCGACAGCAGTGGTGGCGACAGCCCGGAGCGCTGCGGCTTCTTCGAGATCTC 530
 186 GCGGAGATCGACAGCAGTGGTGGCGACAGCCCGGAGCGCTGCGGCTTCTTCGAGATCTC 245
 531 GCGGAGATCGACAGCAGTGGTGGCGACAGCCCGGAGCGCTTCTTCGAGATCTC 590
 246 GCGGAGATCGACAGCAGTGGTGGCGACAGCCCGGAGCGCTTCTTCGAGATCTC 305
 591 GCGGAGATCGACAGCAGTGGTGGCGACAGCCCGGAGCGCTTCTTCGAGATCTC 650
 306 GCGGAGATCGACAGCAGTGGTGGCGACAGCCCGGAGCGCTTCTTCGAGATCTC 365
 651 GCGGAGATCGACAGCAGTGGTGGCGACAGCCCGGAGCGCTTCTTCGAGATCTC 710
 366 GCGGAGATCGACAGCAGTGGTGGCGACAGCCCGGAGCGCTTCTTCGAGATCTC 425
 711 GCGGAGATCGACAGCAGTGGTGGCGACAGCCCGGAGCGCTTCTTCGAGATCTC 770
 426 GCGGAGATCGACAGCAGTGGTGGCGACAGCCCGGAGCGCTTCTTCGAGATCTC 485
 771 GCGGAGATCGACAGCAGTGGTGGCGACAGCCCGGAGCGCTTCTTCGAGATCTC 830
 486 GCGGAGATCGACAGCAGTGGTGGCGACAGCCCGGAGCGCTTCTTCGAGATCTC 545
 831 CTGCGTCATCAGCTAG 846
 546 CTGCGTCATCAGCTAG 561

RESULT 11

BI596688

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BI596688 814 bp mRNA linear EST 07-SEP-2001
 603243262F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5285894 5',
 mRNA sequence.
 BI596688
 BI596688.1 GI:15489627
 EST.
 human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 814)
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgaops-remail.nih.gov
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
 Toshiyuki and Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM1722 row: k column: 15
 High quality sequence stop: 803.

FEATURES

Source

1. .814

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:5285894"

/clone_lib="NIH_MGC_96"

/tissue_type="hypothalamus"

/lab_host="DH10B"

541 AACAGCAGCTGG-ACAGATGTTCCGCGCTCTTCG-CCATGGCCAAAGCTGCCAG-C 597
 741 AACAGCAGCTGGNACCAGATGTTCCGCGGCTCTTCGCGCCATGGCCAGCTGCCAGCC 800
 598 GAGATGAG-CCAGACCTGACACC-CAAGGTCTCGTTCAGT-----ACTGCCAGCT 647
 801 GAGATGAGCCCGCCAGCTGGACCGCAAGGTCTCGGTGCAATTACATGCAACGGGCTGC 860
 648 GCTGCACAAAGAGCGCTGCGGACAGACAGTCTCTCGGCGCGGAGCGCGCGCGCG 707
 861 ACAAAAAGCGCTGCGGGAACAAACCTTCTGCGGCGCGCGGCGCGCAACCGCGCGC 920
 708 CGCGGACCGCGCGAGCTTGGCA 733
 921 GGGCGGCGCGCGCAACACCGGGGA 946

RESULT 10

BI553776

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BI553776 758 bp mRNA linear EST 05-SEP-2001
 603190722F1 NIH_MGC_95 Homo sapiens cDNA clone IMAGE:5262101 5',
 mRNA sequence.
 BI553776
 BI553776.1 GI:15441088
 EST.
 human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 758)
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgaops-remail.nih.gov
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
 Toshiyuki and Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM1660 row: 1 column: 06
 High quality sequence stop: 738.

FEATURES

source

1. .758

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:5262101"

/clone_lib="NIH_MGC_95"

/tissue_type="hippocampus"

/lab_host="DH10B"

/note="Organ: brain; Vector: pBluescriptR (modified
 pBluescript KS+); Site: 1: BamHI; Site 2: SalI-XhoI (gtcagag
); Oligo-dr primed using primer 5'-TTTTTTTGTGTTT-3',
 size-selected for average insert size 2.5 kb and
 normalized to 10^5. This is a primary library enriched
 for full-length clones and constructed using the
 Cap-trapper method (Carninci, in preparation). Library
 constructed by M. Brownstein (NIH/NHGRI, National
 Institutes of Health). Note: this is a NIH_MGC Library."

BASE COUNT

ORIGIN

Query Match 63.9%; Score 540.8; DB 13; Length 758;

Best Local Similarity 99.5%; Pred. No. 2.4e-96;

Matches 553; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

292 GTTTTCATCTGGTTCAGTCTGGACACCGCGCTCTCGAGAGGTGAGCGGCTC 351

6 GTTCTCATCTGGTTCAGTCTGGACACCGCGGCTCTTCTGAGGAGGTGAGCGGCTC 65

/note="Organ: brain; Vector: pBluescriptR (modified pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcagag); Oligo-dr primed using primer 5'-TTTTTTTTTTTNN-3', size-selected for average insert size 2.3 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIMH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."

BASE COUNT 160 a 310 c 225 g 119 t

ORIGIN

Query Match 63.4%; Score 536.2; DB 13; Length 814;
Best Local Similarity 98.0%; Pred. No. 1.9e-95;
Matches 585; Conservative 0; Mismatches 8; Indels 4; Gaps 4;

QY 1 ATGAACCTGGCGCGATGATCAAGAAGATGTGCCGAGCGACTCGGAGCTGAGTATCCCG 60
Db 220 ATGAACCTGGCGCGATGATCAAGAAGATGTGCCGAGCGACTCGGAGCTGAGTATCCCG 60

QY 61 GCCAAGAACTGCTATCGCATGGTCAATCGCTGCTCCAGGTGGGCAAGAGCGGCATC 120
Db 280 GCCAAGAACTGCTATCGCATGGTCAATCGCTGCTCCAGGTGGGCAAGAGCGGCATC 120

QY 121 GTGTGCGGCTTCTACCGCGCCCTTCGAGAGCGCTTACACGCTTACCATCGAGACTTC 180
Db 340 GTGTGCGGCTTCTACCGCGCCCTTCGAGAGCGCTTACACGCTTACCATCGAGACTTC 180

QY 181 CACCGCAAGTCTACTCCATCCGCGGAGGCTTACACGCTTACCATCGAGACTTC 240
Db 400 CACCGCAAGTCTACTCCATCCGCGGAGGCTTACACGCTTACCATCGAGACTTC 240

QY 241 GGCAACACCGCTTCCCGCCATGCGCGGCTTCCATCTCCTCAGAGAGGCTGACGCTTC 300
Db 460 GGCAACACCGCTTCCCGCCATGCGCGGCTTCCATCTCCTCAGAGAGGCTGACGCTTC 300

QY 301 CTGGTGTTCAGTCTGACACCGCGACTCCTTCGAGAGGCTGACGCTTCAGGAGCAG 360
Db 520 CTGGTGTTCAGTCTGACACCGCGACTCCTTCGAGAGGCTGACGCTTCAGGAGCAG 360

QY 361 ATCTCGACACCAAGTCTTCCCTCAAGAACAAACCAAGGAGAGCTGACGCTTCAGG 420
Db 580 ATCTCGACACCAAGTCTTCCCTCAAGAACAAACCAAGGAGAGCTGACGCTTCAGG 420

QY 421 GTCATCTGCGGCAACAGGCTGACCGGCTTACCGCGAGGCTGACCGAGGAGATC 480
Db 640 GTCATCTGCGGCAACAGGCTGACCGGCTTACCGCGAGGCTGACCGAGGAGATC 480

QY 481 GAGCAGCTGGT-GGGCGAGCAGCCCGGCTGCTTACCGCGAGGCTGACCGAGGAGATC 538
Db 700 GAGCAGCTGGTGGGCGAGAGCCCGGCTGCTTACCGCGAGGCTGACCGAGGAGATC 538

QY 539 AGAAGCAGCAGCTGACACAGATGTTCCGCGGCTTTCGCGCATGCGCAAGCTGCGCA 595
Db 760 AGAAGCAGCAGCTGACACAGATGTTCCGCGGCTTTCGCGCATGCGCAAGCTGCGCA 595

RESULT 12
BM921656
LOCUS
DEFINITION AGENCOURT_6708111 NIH_MGC_115 Homo sapiens linear EST 12-MAR-2002
5', mRNA sequence.
ACCESSION BM921656
VERSION BM921656
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1137)
AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.
Email: cgabbs-re@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM12788 Row: 1 Column: 23
High quality sequence stop: 542.

FEATURES

Location/Qualifiers
1..1137
source

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5753182"
/clone_lib="NIH_MGC_115"
/lab_host="DH10B"
/note="Organ: pooled brain, lung, testis; Vector: PCW-SF0R6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27; 1 male lung, age 27; and 1 male testis, age 69. Library is oligo-dr primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.8 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 021. Note: this is a NIH_MGC Library."

BASE COUNT 223 a 427 c 315 g 172 t

ORIGIN

Query Match 61.6%; Score 521.4; DB 14; Length 1137;
Best Local Similarity 94.3%; Pred. No. 1.6e-92;
Matches 574; Conservative 0; Mismatches 31; Indels 4; Gaps 3;

QY 1 ATGAACCTGGCGCGATGATCAAGAAGATGTGCCGAGCGACTCGGAGCTGAGTATCCCG 60
Db 221 ATGAACCTGGCGCGATGATCAAGAAGATGTGCCGAGCGACTCGGAGCTGAGTATCCCG 60

QY 61 GCCAAGAACTGCTATCGCATGGTCAATCGCTGCTCCAGGTGGGCAAGAGCGGCATC 120
Db 281 GCCAAGAACTGCTATCGCATGGTCAATCGCTGCTCCAGGTGGGCAAGAGCGGCATC 120

QY 121 GTGTGCGGCTTCTCACCAGCGGCTTCGAGAGCGCTTACCGAGCTGACATCTCGACAGCTTC 180
Db 341 GTGTGCGGCTTCTCACCAGCGGCTTCGAGAGCGCTTACCGAGCTGACATCTCGACAGCTTC 180

QY 181 CACCGCAAGTCTTACTCCATCCGCGGAGGCTTACCGAGCTGACATCTCGACAGCTTC 400
Db 401 CACCGCAAGTCTTACTCCATCCGCGGAGGCTTACCGAGCTGACATCTCGACAGCTTC 400

QY 241 GGCAACACCGCTTCCCGCCATGCGCGGCTTCCATCTCCTCAGAGAGGCTGACGCGCTCAGGAGCAG 460
Db 461 GGCAACACCGCTTCCCGCCATGCGCGGCTTCCATCTCCTCAGAGAGGCTGACGCGCTCAGGAGCAG 460

QY 301 CTGTGTTCAGTCTGACACCGCGCTTCCATCTCCTCAGAGAGGCTGACGCGCTCAGGAGCAG 520
Db 521 CTGTGTTCAGTCTGACACCGCGCTTCCATCTCCTCAGAGAGGCTGACGCGCTCAGGAGCAG 520

QY 361 ATCTCGACACCAAGTCTTCCCTCAAGAACAAACCAAGAGAGCTGAGAGCTGCGCCCTG 420
Db 581 ATCTCGACACCAAGTCTTCCCTCAAGAACAAACCAAGAGAGCTGAGAGCTGCGCCCTG 420

QY 421 GTCATCTGCGGCAACAGGCTGACCGGCTTCTACCGGAGGTGACGCGCTCAGGAGCAG 640
Db 641 GTCATCTGCGGCAACAGGCTGACCGGCTTCTACCGGAGGTGACGCGCTCAGGAGCAG 640

QY 480 CGAGCAGCTGGTGGGCGAGCAGCCCGGCTTCTCGAGAGCTGCGGCTTCTCGAGATCTCGGCAAGAA 539
Db 701 CGAGCAGCTGGTGGGCGAGCAGCCCGGCTTCTCGAGAGCTGCGGCTTCTCGAGATCTCGGCAAGAA 539

QY 540 GACACAGCAGCTTGGACACAGATGTTCCGCGGCTTCTCGGCGCTGCGCAAGCTGCGCCAGC 597
Db 540 GACACAGCAGCTTGGACACAGATGTTCCGCGGCTTCTCGGCGCTGCGCAAGCTGCGCCAGC 597

Db	760	AAACAGCATCCCTGGACCCCAATGTTCCGGGGCTCTCTCCCTCCCTGGCCCAACCTGCCAG	819
Qy	598	GAGATGAGC	606
Db	820	CGGAGAGC	828
RESULT 13			
BM805574		1160 bp	linear
AGENCY	AGENCY	NIH_MGC_124	Homo sapiens
DEFINITION	5', mRNA sequence.		
ACCESSION	BM805574		
VERSION	BM805574.1	GI:19122397	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 1160)		
TITLE	NIH-MGC http://mgc.nci.nih.gov/.		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: Invitrogen CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM12725 row: o column: 18 High quality sequence stop: 23 High quality sequence stop: 501.		
FEATURES	Location/Qualifiers		
source	1..1160		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:5729129"		
	/clone_lib="NIH_MGC_124"		
	/tissue_type="hippocampus"		
	/lab_host="DH10B"		
	/note="Organ: brain; Vector: pCMV-Sport6; Site: 1: EcoRV (destroyed); Site: 2: NotI; RNA source male hippocampus, age 27. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 0.9-4 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 012."		
BASE COUNT	207 a 473 c 279 g 201 t		
ORIGIN			
Query Match	61.68; Score 521; DB 14; Length 1160;		
Best Local Similarity	97.18; Pred. No. 1.9e-92;		
Matches	541; Conservative 0; Mismatches 15; Indels 1; Gaps 1;		
Qy	1	ATGAAGTGGCGGATGATCAAGAAGATGTGCCCGGAGCTCGGAGCTGAGTATCCCG	60
Db	197	ATGAAGTGGCGGATGATCAAGAAGATGTGCCCGGAGCTCGGAGCTGAGTATCCCG	256
Qy	61	GCAAGAACTGTATCGATGATGATCTCGGCTCGTCCAGGTGGCGAGCGCCATC	120
Db	257	GCAAGAACTGTATCGATGATGATCTCGGCTCGTCCAGGTGGCGAGCGCCATC	316
Qy	121	GTGTCGGCTCTCCTACCGCGGCTTCGAGGAGCGCTACAGCTACAGCTACAGTTC	180
Db	317	GTGTCGGCTCTCCTACCGCGGCTTCGAGGAGCGCTACAGCTACAGTTC	376
Qy	181	CACCGCAAGTTTACTTCCATCCGGGGAGGTCTACAGCTCGACATCTCGACAGTCC	240
Db	377	CACCGCAAGTTTACTTCCATCCGGGGAGGTCTACAGCTCGACATCTCGACAGTCC	436
Qy	241	GGCAACACCCGTTCCCGCATGCGCGCTCTCCATCTCACAGGAGACGCTTTTCATC	300
Db	437	GGCAACACCCGTTCCCGCATGCGCGCTCTCCATCTCACAGGAGACGCTTTTCATC	496
Qy	301	CTGGTGTTCAGTCTGGACAAACCGCGACTCTTCGAGGAGTTCGAGCGCTCAGGCGAGC	360
Db	497	CTGGTGTTCAGTCTGGACAAACCGCGACTCTTCGAGGAGTTCGAGCGCTCAGGCGAGC	556
Qy	361	ATCCTGACACCACTCTTCCTCAAGAAACCAAGGAGAACCTGGACCTGCCCCCTG	420
Db	557	ATCCTGACACCACTCTTCCTCAAGAAACCAAGGAGAACCTGGACCTGCCCCCTG	616
Qy	421	GTCAATCTCGCGCAACAAAGGTGACCGCGACTTCTACCGG-AGTGGACAGCGCGAGT	479
Db	617	GCCATCTCGCGCAACAAAGGTGACCGCGACTTCTACCGGAGGTGGACAGCGCGAGT	676
Qy	480	CGAGCAGCTGTGGCGGACGACCCCGCTGCGCTTCTCGAGATCTCGGCCCAAGAA	539
Db	677	CGAGCAGCTGTGGCGGACGACCCCGCTGCGCTTCTCGAGATCTCGGCCCAAGAA	736
Qy	540	GAACAGCAGCTGGACCC 556	
Db	737	AAACACCCCTGGAGC 753	
RESULT 14			
BM805637		831 bp	mRNA
LOCUS	603243203F1 NIH_MGC_96	Homo sapiens	linear
DEFINITION	mRNA sequence.		EST 07-SEP-2001
ACCESSION	BM805637		
VERSION	BM805637.1	GI:15489576	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 831)		
TITLE	NIH-MGC http://mgc.nci.nih.gov/.		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: Invitrogen CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM1722 row: e column: 01 High quality sequence stop: 777.		
FEATURES	Location/Qualifiers		
source	1..831		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:5285736"		
	/clone_lib="NIH_MGC_96"		
	/tissue_type="hypothalamus"		
	/lab_host="DH10B"		
	/note="Organ: brain; Vector: pBluescriptR (modified pBluescript KS+); Site: 1: BamHI; Site: 2: SalI-XhoI (gtcgag); Oligo-dT primed using primer 5'-TTTTTTTATTTTATTTN-3', size-selected for average insert size 2.3 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIMH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."		
BASE COUNT	167 a 311 c 233 g 120 t		
ORIGIN			

Query Match 61.4%; Score 519.6; DB 13; Length 831;
 Best Local Similarity 96.7%; Pred. No. 3.5e-92;
 Matches 584; Conservative 0; Mismatches 14; Indels 6; Gaps 5;

QY 12 CGCGATGATCAAGAAGATGTCGCCGAGCGACTCGGAGCTGAGTATCCCGGCGCAAGAACTG 71
 Db 230 CGCGATGATCAAGAAGATGTCGCCGAGCGACTCGGAGCTGAGTATCCCGGCGCAAGAACTG 289

QY 72 CTATCCGATGGTATCTCTCGGCTCTGCAAGTGGGCAAGCGGCATCGTGTGCGCGTT 131
 Db 290 CTATCCGATGGTATCTCTCGGCTCTGCAAGTGGGCAAGCGGCATCGTGTGCGCGTT 349

QY 132 CTTACCGCGCGCTTCGAGGAGCTTACAGCTCGACATCTCTGACACGTCGCGCAACACC 251
 Db 350 CTTACCGCGCGCTTCGAGGAGCTTACAGCTCGACATCTCTGACACGTCGCGCAACACC 409

QY 192 CTACTCCATCCCGCGCGCTTCGAGGAGCTTACAGCTCGACATCTCTGACACGTCGCGCAACACC 251
 Db 410 CTACTCCATCCCGCGCGCTTCGAGGAGCTTACAGCTCGACATCTCTGACACGTCGCGCAACACC 469

QY 252 GTTCCCGCGCATCGCGCGCTTCGAGGAGCTTACAGCTCGACATCTCTGACACGTCGCGCAACACC 311
 Db 470 GTTCCCGCGCATCGCGCGCTTCGAGGAGCTTACAGCTCGACATCTCTGACACGTCGCGCAACACC 529

QY 312 TCTGACAAACCGCGACTCTCTCGAGGAGTGCAGCGCTCAGGAGAGTCTCTGACATCTCTGACAC 371
 Db 530 TCTGACAAACCGCGACTCTCTCGAGGAGTGCAGCGCTCAGGAGAGTCTCTGACATCTCTGACAC 589

QY 372 CAAGTCTTGGCTCAAGAAACAAACCAAGGAGAGCTGAGAGTGCAGCGCTCTGACATCTCTGCGG 431
 Db 590 CAAGTCTTGGCTCAAGAAACAAACCAAGGAGAGCTGAGAGTGCAGCGCTCTGACATCTCTGCGG 649

QY 432 CAACAAGGAGTGCAGCGCTCTACCGCGA-GTGGAGCAGCGGA-GATCGAGCAGCT 488
 Db 650 CAACAAGGAGTGCAGCGCTCTACCGCGAGGAGTGGAGCAGCGGAAGATCGAGCAGCT 709

QY 489 GGTGGGAGACACCCCGCGCTCTGAGAGTGCAGCGCTCTGAGAGTCTGCGCAAGAAAGACAGCAG 548
 Db 710 GGTGGGAGACACCCCGCGCTCTGAGAGTGCAGCGCTCTGAGAGTCTGCGCAAGAAAGACAGCAG 548

QY 549 CTTGACAGATGTCGCGCGCTCTGCGCATGCGGAGCTGCGGAGATCGCG-CAAGAAGACAGCAG 608
 Db 768 CTTGACAGATGTCGCGCGCTCTGCGCATGCGGAGCTGCGGAGATCGCG-CAAGAAGATGAGCCC 826

QY 609 AGAC 612
 Db 827 AGAC 830

RESULT 15
 BG711792
 LOCUS
 DEFINITION
 655 bp mRNA linear EST 08-MAY-2001
 pglin.pk009.i15 Normalized Liver Library gallus cdna clone
 pglin.pk009.i15 5' similar to gb|AAAF3090.1|AF239157_1 (AF239157)
 DEXR51 [Rattus norvegicus]G, mRNA sequence.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 BG711792.1 GI:14005742
 EST.
 chicken.
 Gallus gallus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianinae; Gallus.
 1 (bases 1 to 655)
 Burnside, J., Morgan, R.W. and Cogburn, L.A.
 Chicken ESTs from a normalized liver library
 Unpublished (2001)
 Contact: Joan Burnside
 Molecular Endocrinology
 University of Delaware
 40 Townsend Hall, Newark, DE 19717, USA
 Tel: 302 831-1345

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

FEATURES
 source
 1..655
 /organism="Gallus gallus"
 /db_xref="taxon:9031"
 /clone="pglin.pk009.i15"
 /clone_lib="Normalized Liver Library"
 /sex="Male and Female"
 /tissue_type="liver"
 /lab_host="E.coli EMDH10B"
 /note="Vector: pCMVSPORT 6"

BASE COUNT 145 a 206 c 184 g 104 t 16 others
 ORIGIN

Query Match 60.7%; Score 513.6; DB 12; Length 655;
 Best Local Similarity 88.3%; Pred. No. 5e-91;
 Matches 560; Conservative 0; Mismatches 73; Indels 1; Gaps 1;

QY 1 ATGAACCTGGCGCGATGATCAAGAAGATGTCGCCGAGCGACTCGGAGCTGAGTATCCCG 60
 Db 14 ATGAACCTGGCGCGATGATCAAGAAGATGTCGCCGAGCGACTCGGAGCTGAGTATCCCG 73

QY 61 GCCAAGAACTGCTATCGCATGCTCATCTCGGCTCTGAGGAGCGCTTACACGCTACCATCGAGGACTTC 120
 Db 74 GCCAAGAACTGCTATCGCATGCTCATCTCGGCTCTGAGGAGCGCTTACACGCTACCATCGAGGACTTC 133

QY 121 GTGTCGCGCTTCTCACCAGCGGCTTTCGAGGAGCGCTTACACGCTACCATCGAGGACTTC 180
 Db 134 GTGTCGCGCTTCTCACCAGCGGCTTTCGAGGAGCGCTTACACGCTACCATCGAGGACTTC 193

QY 181 CACGCAAGTTCTACTCCATCCGCGCGAGGTCTACACGCTCGACATCTCTGAGGAGCTTC 240
 Db 194 CACGCAAGTTCTACTACGATCCGCGGTGAGGTCTACACGCTCGACATCTCTGAGGAGCTTC 253

QY 241 GGCAACCAACCGCTTCCCGCGCTTCCGCGCTCTCCATCTCTCAGGAGAGCTTTTCATC 300
 Db 254 GGCAACCAACCGCTTCCCGCGCTTCCGCGCTTCCGCGCTCTCCATCTCTCAGGAGAGCTTTTCATC 313

QY 301 CTGCTGTTCAGTCTGGACAACCGCGACTCTCTGAGGAGGTGCGAGGCTCAGGCGCTCAGGAGCAG 360
 Db 314 CTGCTGTTCAGTCTGGACAACCGCGACTCTCTGAGGAGGTGCGAGGCTCAGGCGCTCAGGAGCAG 373

QY 361 ATCTCTGACACCAAGTCTTGCCTCAAGAACAAACCAAGGAGAGCTGGACGTCGCCCTG 420
 Db 374 ATCTCTGACACCAAGTCTTGCCTCAAGAACAAACCAAGGAGAGCTGGACGTCGCCCTG 433

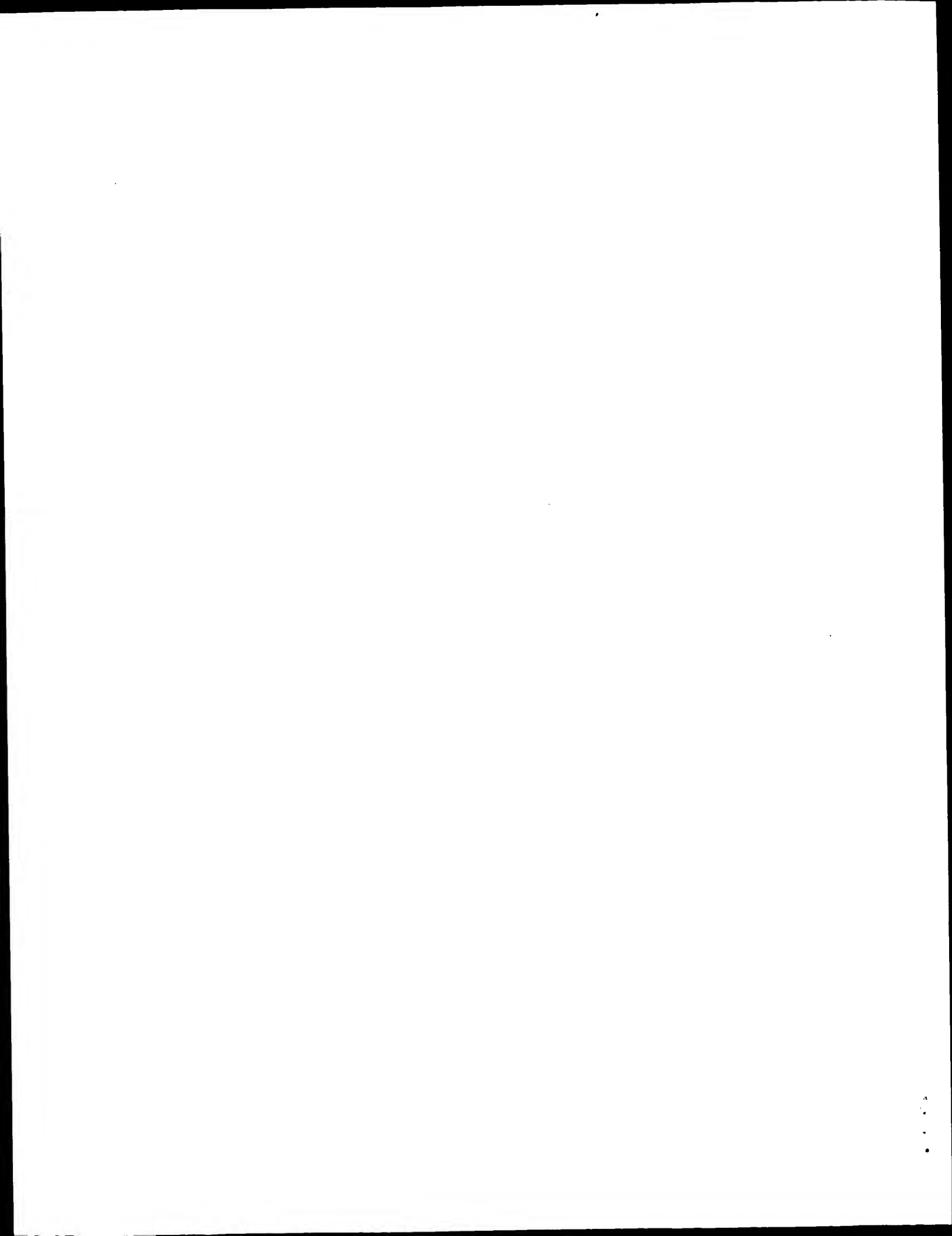
QY 421 GTCATCTCGCGCAACAGGAGTGCAGCGGACTTCTACCGGAGGTGGACGCGGAGATC 480
 Db 434 GTCATCTCGCGCAACAGGAGTGCAGCGGACTTCTACCGGAGGTGGACGCGGAGATC 493

QY 481 GAGCAGCTGGTGGCGACACCGCGCGCTGCGC-CTACTTCGAGATCTCGGCAAGNA 539
 Db 494 GAGCAGCTGGTGGCGACAGACCGCGCAAGAAATGCGGNNNACTTCGAGATCTCGGCAAGNA 553

QY 540 GAACAGCAGCTGGACAGATGTTCCGCGGCTTTCGCCATGGCCAGCTGCCCGGGA 599
 Db 554 GAACAGCAGCTGGATGATGTTCCAGGCGCTTTCGCCATGGCCAGCTGCCCGGGA 613

QY 600 GATGAGCCAGACTGCACCGCAAGGTCTCGGTG 633
 Db 614 GATGAGCCAGACTGCACCGGNNNNNNNNNNNG 647

Search completed: December 28, 2002, 05:50:39
 Job time : 1127.58 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 28, 2002, 02:13:37 ; Search time 4571.56 Seconds
(without alignments)
11465.267 Million cell updates/sec

Title: US-09-709-103-3

Perfect score: 1801

Sequence: 1 ggaattccgagcgagccgg.....toatgtaattatgaac 1801

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sy.*

13: gb.un.*

14: gb.vi.*

15: em.ba.*

16: em.fun.*

17: em.hum.*

18: em.in.*

19: em.mu.*

20: em.om.*

21: em.or.*

22: em.ov.*

23: em.pat.*

24: em.ph.*

25: em.pl.*

26: em.ro.*

27: em.sts.*

28: em.un.*

29: em.vi.*

30: em.htg.hum.*

31: em.htg.inv.*

32: em.htg.other.*

33: em.htg.mus.*

34: em.htg.pln.*

35: em.htg.rod.*

36: em.htg.mam.*

37: em.htg.vrt.*

38: em.sy.*

39: em.htgo.hum.*

40: em.htgo.mus.*

41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1740	96.6	1740	9	AF069506 Homo sapi
2	1691.4	93.9	1758	9	BC018041 Homo sapi
3	1673.4	92.9	1746	9	AF177335 Homo sapi
4	1460	81.1	4990	9	AF222979 Homo sapi
5	1458.4	81.0	18334	9	AC020558 Homo sapi
6	1431.2	79.5	5141	9	AF262018 Homo sapi
7	1430.2	79.4	58882	2	AC073621 Homo sapi
8	1389.8	77.2	183598	2	AC090608 Homo sapi
9	1137	63.1	1187	9	AF153192 Homo sapi
10	940.4	52.2	979	9	AF172846 Homo sapi
11	846	47.0	846	9	AF498923 Homo sapi
12	764.8	42.5	1612	10	BC034166 Mus muscu
13	764.2	42.4	1623	10	AF009246 Mus muscu
14	746.8	41.5	1616	10	AF239157 Rattus no
15	592.2	32.9	179124	10	AL603710 Mouse DNA
16	468.2	26.0	162504	2	AC122995 Rattus no
17	411.4	22.8	2973	6	AX393267 Sequence
18	411.4	22.8	3058	9	BC013419 Homo sapi
19	407.4	22.6	2832	6	AX393244 Sequence
20	407.4	22.6	2832	9	AF279143 Homo sapi
21	362.2	20.1	2699	9	HSB03172 Homo sapi
22	361	20.0	3469	10	AF134409 Rattus no
23	359.4	20.0	3020	6	AX393362 Sequence
24	352.6	19.6	414	6	AX336014 Sequence
25	352.6	19.6	414	6	AX336237 Sequence
26	352.6	19.6	414	6	AX336666 Sequence
27	352.6	19.6	414	6	AX410244 Sequence
28	240.2	13.3	114771	9	HS569D19 Human DNA
29	205.8	11.4	211071	10	AC076974 Mus muscu
30	181.8	10.1	66031	2	AC118046 Mus muscu
31	136.2	7.6	774	9	AB076888 Homo sapi
32	136.2	7.6	2827	9	AK096600 Homo sapi
33	136.2	7.6	3391	9	BC030660 Homo sapi
34	136.2	7.6	177540	9	AC006538 Homo sapi
35	134.8	7.5	597	9	AY056037 Homo sapi
36	134.8	7.5	597	9	AY059641 Homo sapi
37	132	7.3	3061	6	AX430418 Sequence
38	129	7.2	2294	10	BC026377 Mus muscu
39	125.8	7.0	1249	6	AX430295 Sequence
40	116	6.4	4412	9	AB062937 Macaca fa
41	114.8	6.4	1913	9	BC012362 Homo sapi
42	114.8	6.4	140794	2	AC069415 Homo sapi
43	114.8	6.4	165733	2	AC069072 Homo sapi
44	114.8	6.4	198170	9	AC117394 Homo sapi
45	113.6	6.3	2145	9	HSB03027 Homo sapi

ALIGNMENTS

RESULT 1
AF069506
LOCUS
DEFINITION Homo sapiens activator of G protein signaling (AGS1) mRNA, complete
AF069506
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS

AF069506
Homo sapiens
cds.
AF069506
Homo sapiens
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1740)
Cismowski, M.J., Takesono, A., Ma, C., Lizano, J.S., Xie, X.,
Fuernkranz, H., Lanier, S.M. and Duzic, E.

AF069506
Homo sapiens
cds.
AF069506
Homo sapiens
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1740)
Cismowski, M.J., Takesono, A., Ma, C., Lizano, J.S., Xie, X.,
Fuernkranz, H., Lanier, S.M. and Duzic, E.

TITLE	Genetic screens in yeast to identify mammalian nonreceptor modulators of G-protein signaling
JOURNAL	Nat. Biotechnol. 17 (9), 878-883 (1999)
MEDLINE	99403338
PUBMED	10471929
REFERENCE	2 (bases 1 to 1740)
AUTHORS	Cismowski,M.J., Fuernkrantz,H., Ma,C., Spruyt,M., Xie,X., Lanier,S.M. and Duzic,E.
TITLE	Direct Submission
JOURNAL	Submitted (01-JUN-1998) Receptor Pharmacology/Biochemistry, Cadus Pharmaceutical Corporation, 777 Old Saw Mill River Rd., Tarrytown, NY 10591, USA
FEATURES	<p>Location/Qualifiers</p> <p>1..1740</p> <p>/organism="Homo sapiens"</p> <p>/db_xref="taxon:9606"</p> <p>/tissue_type="liver"</p> <p>/dev_stage="adult"</p> <p>1..1740</p> <p>/gene="AGS1"</p> <p>146..991</p> <p>/gene="AGS1"</p> <p>/note="ras-related G protein; similar to Mus musculus product encoded by GenBank Accession Number AF009246"</p> <p>/codon_start=1</p> <p>/product="activator of G protein signaling"</p> <p>/protein_id="AAD34206.1"</p> <p>/db_xref="GI:4959038"</p> <p>/translation="MKLAAMTKMKPDSSELSIPAKNCYRMVILGSKVKGKTAIVS LGRFEDATPTIEDPHRFKFSYIRGEVYQDLDTSGNHPPAMRRRLSILTDGVFIIF FSLDNKSEFEEVQRRLQILDTKSLKNKTKENVDVPLVICGNKGRDRDFYREVDOR EQLVGDDPORCAFEISAKNSKSLDOMFRALFNAAKLSEMSPLHRKVSVOYCDVT KKAIRNKKLLRAGSGGGGGPDGAFGIAPFARRPSVHSDILMIIREKASAGSQAKD RCVLS"</p>
BASE COUNT	422 a 546 c 483 g 289 t
ORIGIN	
Query Match	96.6%; Score 1740; DB 9; Length 1740;
Best Local Similarity	100.0%; Pred. No. 1.3e-219;
Matches 1740; Conservative	0; Mismatches 0; Indels 0; Gaps
QY	9 GAGCGGAGCCGGAGCCCAAGCCGAGCCGCGCCGACCGCCGAGCAGAGCCCTCCAGCCGC 68
DB	1 GAGCGGAGCCGGAGCCCAAGCCGAGCCGCGCCGACCGCCGAGCAGAGCCCTCCAGCCGC 60
QY	69 TCACCCCGCGTGCACACCCAGCAGACCTCAGCGCTCTCTGCGCTTCTCTGCGCCCGCG 128
DB	61 TCACCCCGCGTGCACACCCAGCAGACCTCAGCGCTCTCTGCGCTTCTCTGCGCCCGCG 120
QY	129 CCGCGCTTCGCGCCCTCTGCCCAATGAACCTGGCGCGATGATCAAGAAGATGTGCC 188
DB	121 CCGCGCTTCGCGCCCTCTGCCCAATGAACCTGGCGCGATGATCAAGAAGATGTGCC 180
QY	189 GAGCGACTCGGAGCTGAGTATCCGCGCCAGAAGCTGATTCGATGGTGTATCTCGGCTC 248
DB	181 GAGCGACTCGGAGCTGAGTATCCGCGCCAGAAGCTGATTCGATGGTGTATCTCGGCTC 240
QY	249 GTCCAAAGTGGCCAGACGCGCATCGTGCAGCTTCTCACCGCGCGCTTCGAGGACGC 308
DB	241 GTCCAAAGTGGCCAGACGCGCATCGTGCAGCTTCTCACCGCGCGCTTCGAGGACGC 300
QY	309 CTACAGCGCTACCATCGAGGACTTCCACCGCAAGTTCTACTCCATCCGCGCGAGGCTTA 368
DB	301 CTACAGCGCTACCATCGAGGACTTCCACCGCAAGTTCTACTCCATCCGCGCGAGGCTTA 360
QY	369 CCAGCTCGACATCTCTGACACGTCGCGCAACACCGCTTCCCGCCATCGGCGCTCTC 428
DB	361 CCAGCTCGACATCTCTGACACGTCGCGCAACACCGCTTCCCGCCATCGGCGCTCTC 420
QY	429 CATCTTCACAGGAGAGCTTTTCATCTCGTGTTTCAGTCTGGACACACCGGACTCCTTCGA 488
DB	421 CATCTTCACAGGAGAGCTTTTCATCTCGTGTTTCAGTCTGGACACACCGGACTCCTTCGA 480


```

|||||
Db 1561 GTCACATGGACGACAAACCTTACCAGGTGTTTACTGTGTGTGTGAGGTCT 1620
Qy 1629 TTAAGTTATGCTTTATTTGGTTTTTAAATATACATAAATAATTTAAATGGAATA 1688
Db 1621 TTAAGTTATGCTTTATTTGGTTTTTAAATATACATAAATAATTTAAATGGAATA 1680
Qy 1669 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1748
Db 1681 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1740

RESULT 2
BC018041
LOCUS BC018041
DEFINITION Homo sapiens, RAS, dexamethasone-induced 1, clone MGC:26290
ACCESSION BC018041
VERSION BC018041.1
KEYWORDS GI:17390075
SOURCE MGC.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1758)
Strausberg, R.
Direct Submission
Submitted (03-DEC-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabbs@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki
Toshinuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www.shgc.stanford.edu
Contact: (Dickson, Mark) mcdpaxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAP Plate: 32 Row: f Column: 11.
Location/Qualifiers
1..1758
/organism="Homo sapiens"
/db_xref="LocusID:51655"
/db_xref="taxon:9606"
/clone="MGC:26290 IMAGE:4797482"
/tissue_type="Brain, hypothalamus"
/clone_lib="NIH_MGC_96"
/lab_host="DH10B"
/notes="Vector: pBluescript"
213..1058
/codon_start=1
/product="RAS, dexamethasone-induced 1"
/protein_id="AAH18041.1"
/db_xref="GI:17390076"
/translation="MKLAAMIKMCPDSELSIPAKNRYRMVILGSKVKGKATVSRF
LTGRFEDAVPTIEDPHRFYSIRGEVYQLDILDTSGNHPFAMRRSLITGDVFLV
FSLDNRSFEVRLRQILDTKSLKNKTRNDVDVPLVICGNKGRDFFREVDREI
EQLVGDIPORCAVFIISAKNSLDOMEFALFMAKLPSESPDLHRKVSQYCDVLH
KKALRNKLLRAGSGGGGDPDAFIVAPFARRPSVHSDLMYIREKASQSAKDKE
RCVIS"
BASE COUNT 387 a 576 c 501 g 294 t
ORIGIN

Query Match
93.9%; Score 1691.4; DB 9; Length 1758;
|||||

```

```

Best Local Similarity 99.9%; Pred. No. 3.2e-213;
Matches 1692; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 7 CCGAGCGAGCCGCGAGCCCAAGCCGCGAGCCGCGCCAGCCGAGCAGAGCCCTCCAGCC 66
Db 66 CCGAGCGAGCCGCGAGCCCAAGCCGCGAGCCGCGCCAGCCGAGCAGAGCCCTCCAGCC 125
Qy 67 GCTACACCCCGGTGGCCACCCAGCAGCCCTTCAGCCGCTCTCTGCCCTTCTCTCGCCCG 126
Db 126 GCTACACCCCGGTGGCCACCCAGCAGCCCTTCAGCCGCTCTCTGCCCTTCTCTCGCCCG 185
Qy 127 CGCCCGCCCTCGCGCGCCCTCTGCCCAATGAACACTGGCCGCGCATGATCAAGAAGATGTC 186
Db 186 CGCCCGCCCTCGCGCGCCCTCTGCCCAATGAACACTGGCCGCGCATGATCAAGAAGATGTC 245
Qy 187 CCGAGGACTCGGAGCTGAGTATCCCGGCCAAGAACTGCTATCGCATGCTATCTCCGCG 246
Db 246 CCGAGGACTCGGAGCTGAGTATCCCGGCCAAGAACTGCTATCGCATGCTATCTCCGCG 305
Qy 247 TCGTCCAAAGTGGCAAGACGGCCATCGTGTGCGCTTCTCTACCGGCGCTTCGAGGAC 306
Db 306 TCGTCCAAAGTGGCAAGACGGCCATCGTGTGCGCTTCTCTACCGGCGCTTCGAGGAC 365
Qy 307 GCCTACACGCTTACCATCGAGGACTTCCACCGCAAGTGTCTACTCCATCCGCGGAGGTC 366
Db 366 GCCTACACGCTTACCATCGAGGACTTCCACCGCAAGTGTCTACTCCATCCGCGGAGGTC 425
Qy 367 TACGAGCTCGACATCCTCGACACGTCGCGCAACACCGGTCGCCGCGCATCGCGCGCTC 426
Db 426 TACGAGCTCGACATCCTCGACACGTCGCGCAACACCGGTCGCCGCGCATCGCGCGCTC 485
Qy 427 TCCATCCTCACAGGAGAGCTTTCATCTCTGTTTCTAGTCTGGACAAACCGGACTCTTTC 486
Db 486 TCCATCCTCACAGGAGAGCTTTCATCTCTGTTTCTAGTCTGGACAAACCGGACTCTTTC 545
Qy 487 GAGGAGTGCAGCGGCTCAGGAGCAGATCTCGACACCAAGTCTTCCCTCAGAACAAA 546
Db 546 GAGGAGTGCAGCGGCTCAGGAGCAGATCTCGACACCAAGTCTTCCCTCAGAACAAA 605
Qy 547 ACCAAGGAGAACGTGAGCTGCGCCCTTGGTCTATCTGCGGCAACAGGCTGACCGGACTTC 606
Db 606 ACCAAGGAGAACGTGAGCTGCGCCCTTGGTCTATCTGCGGCAACAGGCTGACCGGACTTC 665
Qy 607 TACGCGAGTGGACCGCGAGATCGAGCTGTGGCGAGACCTGTGGCGAGACCCCGAGCTGC 666
Db 666 TACGCGAGTGGACCGCGAGATCGAGCTGTGGCGAGACCTGTGGCGAGACCCCGAGCTGC 725
Qy 667 GCCTACTTCGAGATCTCGGCCAAGAACAGCAGCCTGGACCAAGATGTTTCCGCGGCTC 726
Db 726 GCCTACTTCGAGATCTCGGCCAAGAACAGCAGCCTGGACCAAGATGTTTCCGCGGCTC 785
Qy 727 TTCGCCATGGCCAAAGCTGCCAGGAGATCGCCGAGATCGCCGAGCTGACCGCAAGTCTCGGTG 786
Db 786 TTCGCCATGGCCAAAGCTGCCAGGAGATCGCCGAGATCGCCGAGCTGACCGCAAGTCTCGGTG 845
Qy 787 CAGTACTGCGAGCTGTGCAACAAGAGGCGCTCGGAACAAGAGCTGCTGCGGCGCGC 846
Db 846 CAGTACTGCGAGCTGTGCAACAAGAGGCGCTCGGAACAAGAGCTGCTGCGGCGCGC 905
Qy 847 AGCGGCGCGCGCGCGAGCCCGCGGAGCTTTGGCATGTGGCAACCTTTCGCGCGC 906
Db 906 AGCGGCGCGCGCGCGAGCCCGCGGAGCTTTGGCATGTGGCAACCTTTCGCGCGC 965
Qy 907 CGGCGCGCGCTACACAGCAGCTTCATGTATCTCGCGAGAGCCAGCGCGCGAGCCAG 966
Db 966 CGGCGCGCGCTACACAGCAGCTTCATGTATCTCGCGAGAGCCAGCGCGCGAGCCAG 1025
Qy 967 GCCAAGGAGAGAGCGCTGCGCTATCAGTATAGGAGCCCGCGCTGGGACACAAACC 1026
Db 1026 GCCAAGGAGAGAGCGCTGCGCTATCAGTATAGGAGCCCGCGCTGGGACACAAACC 1085
Qy 1027 TAAGGAGGAGCTTTTGTAAAGTCAAACTCAACGCGCGGTGCGCCCGGAGGCG 1086
|||||

```

FEATURES		Location/Qualifiers	
source		1. 1746	
		/organism="Homo sapiens"	
		/db_xref="taxon:9606"	
		/clone="SPI942"	
		455..1039	
CDS		/codon_start=1	
		/product="unknown"	
		/protein_id="AAG17979.1"	
		/db_xref="GI:10503969"	
		/translation="MRSLTILTGDNVFLVSLNDRDSFEVQRUQIILDTKSLKNK	
		TKENVPLVTCGNKGNKGRDRFYREVDREIEQLVGDOPRCAYFEISAKNNSUDOMFR	
		ALFAMAKLPSEMSDPLHKKVSVQYCDVHLKALRNKKLLRAGSGGGGDPDAGIVA	
		PFARRSVHSLMYIREKASAGSKDKKCVIS"	
BASE COUNT		388 a 569 c 496 g 293 t	
ORIGIN			
Query Match		92.9%; Score 1673.4; DB 9; Length 1746;	
Best Local Similarity		99.8%; Pred. No. 7.4e-211;	
Matches 1696; Conservative		0; Mismatches 1; Indels 2; Gaps 2;	
QY	7	CCGAGCGGAGCGGAGCGCCCAAGCGCGAGCGCGCGCCCGAGCAGAGCGCCCTCCAGCC	66
Db	46	CCGAGCGGAGCGGAGCGCCCAAGCGCGAGCGCGCGCCCGAGCAGAGCGCCCTCCAGCC	105
QY	67	GCTACACCCCGCGTGCACCCCGAGCGACCTCAGCGCGCTCTCTGCGCTTCTCTCGGCCCG	126
Db	106	GCTACACCCCGCGTGCACCCCGAGCGACCTCAGCGCGCTCTCTGCGCTTCTCTCGGCCCG	165
QY	127	CGCCCGCGCTCGCGCGCCCTCTGCCCAATGAAACTTGGCGCGCATGATCAAGAAGATGTGC	186
Db	166	CGCCCGCGCTCGCGCGCCCTCTGCCCAATGAAACTTGGCGCGCATGATCAAGAAGATGTGC	225
QY	187	CCGAGCGACTCGGAGCTGAGTATCCCGGCGCAAGAACTGCTATCGCATGTGTCATCTCGGC	246
Db	226	CCGAGCGACTCGGAGCTGAGTATCCCGGCGCAAGAACTGCTATCGCATGTGTCATCTCGGC	285
QY	247	TCGTCAAGTGGGCGAGCGCGCATGTCGCGCTTCTCACCAGCGCGCTTCGAGGAC	306
Db	286	TCGTCAAGTGGGCGAGCGCGCATGTCGCGCTTCTCACCAGCGCGCTTCGAGGAC	345
QY	307	GCCTACAGCGCTACCATCGAGGACTTCCACCGCAAGTCTTACTTCCATCGCGGCGAGTTC	366
Db	346	GCCTACAGCGCTACCATCGAGGACTTCCACCGCAAGTCTTACTTCCATCGCGGCGAGTTC	405
QY	367	TACAGC-TGACATCTCGACAGTCCGCAACCGCTTCCCGCGCATGCGCGCGCT	425
Db	406	TACAGCTTTCACATCTCGACAGTCCGCAACCGCTTCCCGCGCATGCGCGCGCT	465
QY	426	CTCCATCTCACAGGAGCGCTTTTCATCTCTGGTGTTCAGTCTGGACAACCGGACTCTT	485
Db	466	CTCCATCTCACAGGAGCGCTTTTCATCTCTGGTGTTCAGTCTGGACAACCGGACTCTT	525
QY	486	CGAGGAGTGCAGGGCTCAGGAGAGATCTTGGCGCAACAAAGGCTGACCGCGACTT	545
Db	526	CGAGGAGTGCAGGGCTCAGGAGAGATCTTGGCGCAACAAAGGCTGACCGCGACTT	585
QY	546	PACCAAGGAGAGCTGGACGTGCCCTTGGTTCATCTGGCGCAACAAAGGCTGACCGCGACTT	605
Db	586	PACCAAGGAGAGCTGGACGTGCCCTTGGTTCATCTGGCGCAACAAAGGCTGACCGCGACTT	645
QY	606	CTACCGGAGTGGGACCGAGCGAGATCGAGTGGTGGGCGACCGCCCGCGCGCTG	665
Db	646	CTACCGGAGTGGGACCGAGCGAGATCGAGTGGTGGGCGACCGCCCGCGCGCTG	705
QY	666	CGCCTACTTCGAGATCTCGGCCAAGAACAGAGCGCTGGACAGATGTTCGCCGCGCT	725
Db	706	CGCCTACTTCGAGATCTCGGCCAAGAACAGAGCGCTGGACAGATGTTCGCCGCGCT	765
QY	726	CTTCGCGATGGCCAAAGCTGCCCGAGAGATGAGCCAGCTGACCGCGAGGCTCGGT	785
Db	766	CTTCGCGATGGCCAAAGCTGCCCGAGAGATGAGCCAGCTGACCGCGAGGCTCGGT	825

Db	1086	TAAGGAGGACCTTTTGTAGTCAAAATCCAAAGCGCGCGGTGCGCCCGAGCGCGGAGCG	1145
QY	1087	CGCGGAGTGCCTCCCTCCCGCGGATCCGCGCCCGAGCTGGGAGCGCCACTG	1146
Db	1146	CGCGGAGTGCCTCCCTCCCGCGGATCCGCGCCCGAGCTGGGAGCGCCACTG	1205
QY	1147	AACCGAAGGAGCGGTCACTGCTCCGAGAGGAAGAGACGGCGGCAAGACTGGGACTA	1206
Db	1206	AACCGAAGGAGCGGTCACTGCTCCGAGAGGAAGAGACGGCGGCAAGACTGGGACTA	1265
QY	1207	TTCCCGACCCCGCTCCCGATGAGCGCGCCACCCCAATCTTGGGAGCGAGGCGC	1266
Db	1266	TTCCCGACCCCGCTCCCGATGAGCGCGCCACCCCAATCTTGGGAGCGAGGCGC	1325
QY	1267	CAGCGAGGCTGATTTATCTTCTCAAGACCTAAGAGTGAAGCGGGGTGGGAGGGA	1326
Db	1326	CAGCGAGGCTGATTTATCTTCTCAAGACCTAAGAGTGAAGCGGGGTGGGAGGGA	1385
QY	1327	TGTGAAGTTATCAGGCTCTGCTAGCTTCAAGAACCGTCAATGCCGCTTGAAGGTTCAG	1386
Db	1386	TGTGAAGTTATCAGGCTCTGCTAGCTTCAAGAACCGTCAATGCCGCTTGAAGGTTCAG	1445
QY	1387	GACCGAGGCTATCTGCTGCTGATTCGCGGTGCTGTGACAGCGGCTAGAGCGCTC	1446
Db	1446	GACCGAGGCTATCTGCTGCTGATTCGCGGTGCTGTGACAGCGGCTAGAGCGCTC	1505
QY	1447	TGCGCTCCGAAACTAAGCGGGGGCGTGGGTCAATCATAGCAAGTGTGTTTAC	1506
Db	1506	TGCGCTCCGAAACTAAGCGGGGGCGTGGGTCAATCATAGCAAGTGTGTTTAC	1565
QY	1507	ATGTGAGTGAACCTGCAAGGAACACAAACAACTTGCCTTAAACGGTAGTTCGG	1566
Db	1566	ATGTGAGTGAACCTGCAAGGAACACAAACAACTTGCCTTAAACGGTAGTTCGG	1625
QY	1567	GTGTCAACATGGACACAAACAACTTACCAGGTGTTTACTGTGTGTGTGAGT	1626
Db	1626	GTGTCAACATGGACACAAACAACTTACCAGGTGTTTACTGTGTGTGTGAGT	1685
QY	1627	CTTTAAAGTTATCTTATTTGTTTTTAAATATACATATAATATAATATAATATAAT	1686
Db	1686	CTTTAAAGTTATCTTATTTGTTTTTAAATATACATATAATATAATATAATATAAT	1745
QY	1687	AAAAAAAAAAAAA 1699	
Db	1746	AAAAAAAAAAAAA 1758	
RESULT 3			
AF177335			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P. Y., Zhao, B., Frengen, E., Tatenoe, M., Catanese, J. J., and de Jong, P. J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pletier de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)

VECTOR: pBAC3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is AC073621. Actual start of this clone is at base position 1 of RP11-524F11; actual end is at base position 183334 of RP11-524F11.

The sequence from position 1987 to 2253 was derived from PCR product of RP11-524F11 BAC DNA.

FEATURES

source		
1. 183334	Location/Qualifiers	
/organism="Homo sapiens"		
/db_xref="taxon:9606"		
/chromosome="17"		
/map="17"		
/clone="RP11-524F11"		
/clone_lib="RPCI-11"		
1. 319		
/note="similar to EST AI672824 (NID:g4852555) w73e06.x1"		
339. .887		
/note="similar to EST AI968384 (NID:g5765202) wu02e05.x1"		
339. .738		
/note="similar to EST AI289781 (NID:g3931622) qw12a07.x1"		
339. .658		
/note="similar to EST AW072137 (NID:g6027135) w298e07.x1"		
340. .767		
/note="similar to EST AI014398 (NID:g3228779) ou93b12.s1"		
340. .686		
/note="similar to EST AW237782 (NID:g6570171) xm81e08.x1"		
346. .742		
/note="similar to EST AI359064 (NID:g4110685) qv25e07.x1"		
405. .884		
/note="similar to EST AI279211 (NID:g3917445) qm24c09.x1"		
782. 1335		
/note="similar to EST BF042464 (NID:g10759519)"		
822. 1319		
/note="similar to EST BE864687 (NID:g10385980)"		
834. .2931		
/note="CpG island (8GC-69.0, o/e=0.80, #CpGs=226)"		
1135. 1585		
/note="similar to EST BF555582 (NID:g11665312)"		
1168. 1583		
/note="similar to EST AW028127 (NID:g5886883) wv26c07.x1"		
1432. 1583		
/note="similar to EST BE390509 (NID:g9335874)"		
1440. 1593		
misc_feature	/note="similar to EST AW915326 (NID:g8081017)"	
1450. 1583		
misc_feature	/note="similar to EST AA790463 (NID:g2850583) vw04g12.r1"	
1795. 2255		
misc_feature	/note="similar to EST BE390509 (NID:g9335874)"	
1795. 2201		
misc_feature	/note="similar to EST AA790463 (NID:g2850583) vw04g12.r1"	
1795. 2153		
misc_feature	/note="similar to EST AW915326 (NID:g8081017)"	
1795. 2114		
misc_feature	/note="similar to EST AW028127 (NID:g5886883) wv26c07.x1"	
1804. 2071		
misc_feature	/note="similar to EST BE836084 (NID:g10268462)"	
1817. 2201		
misc_feature	/note="similar to EST AA672515 (NID:g2644732) vo59e02.r1"	
1844. 2201		
misc_feature	/note="similar to EST AW463433 (NID:g7033601)"	
1985. 2285		
misc_feature	/note="similar to EST AA852097 (NID:g2940690)"	
2538. 3033		
misc_feature	/note="similar to EST BF057610 (NID:g10811506)"	
2554. 2806		
misc_feature	/note="similar to EST BF195653 (NID:g11082770)"	
3561. 3924		
repeat_region	/note="similar to EST R18261 (NID:g771871) yg01a05.r1"	
4027. 4085		
misc_feature	/rpt_family="MER1_type"	
4670. 5117		
misc_feature	/note="similar to EST R42241 (NID:g817108) yf98b09.s1"	
4857. 5117		
repeat_region	/note="similar to EST T16275 (NID:g518437)"	
6155. 6307		
repeat_region	/rpt_family="Alu"	
6309. 6446		
repeat_region	/rpt_family="L2"	
6583. 6891		
repeat_region	/rpt_family="Alu"	
6960. 7153		
repeat_region	/rpt_family="L1"	
7525. 7838		
repeat_region	/rpt_family="Alu"	
7851. 8487		
repeat_region	/rpt_family="L1"	
9000. 9104		
repeat_region	/rpt_family="L2"	
11463. 11750		
misc_feature	/note="similar to EST BF344873 (NID:g11292331)"	
11465. 11750		
misc_feature	/note="similar to EST AI340185 (NID:g4077112) qo79d07.x1"	
11465. 11750		
misc_feature	/note="similar to EST AI571390 (NID:g4534764) tn45e12.x1"	
11465. 11750		
misc_feature	/note="similar to EST AI701138 (NID:g4989038) we10d09.x1"	
11465. 11750		
misc_feature	/note="similar to EST AI803637 (NID:g5369109) tc43d01.x1"	
11465. 11750		
misc_feature	/note="similar to EST BE747895 (NID:g10161887)"	
11465. 11750		
misc_feature	/note="similar to EST BF203429 (NID:g11097015)"	
11465. 11697		
misc_feature	/note="similar to EST AI869572 (NID:g5543540) w198d03.x1"	
11467. 11750		
misc_feature	/note="similar to EST AI817466 (NID:g5436545) wk38g01.x1"	
11467. 11619		
misc_feature	/note="similar to EST AA367144 (NID:g2019483)"	
11468. 11750		
misc_feature	/note="similar to EST BE218333 (NID:g8905651) hv37c09.x1"	
11468. 11750		
misc_feature	/note="similar to EST BE549542 (NID:g9791256)"	
11468. 11750		
misc_feature	/note="similar to EST BE750272 (NID:g10164264)"	
11478. 11750		
misc_feature	/note="similar to EST BF667874 (NID:g11941769)"	

QY 187 CCGAGGAGCTCGAGCTGAGTATCCCGGCCAAGAAGTGTATCGATGCTCATCTCGGC 246
|||||
Db 19093 CCGAGGAGCTCGAGCTGAGTATCCCGGCCAAGAAGTGTATCGATGCTCATCTCGGC 19152
QY 247 TCGTCCAAAGGTGGCAAGACGCGCATCGTGTGCGCGCTTCTCACCAGCGCGCTTCGAGAC 306
|||||
Db 19153 TCGTCCAAAGGTGGCAAGACGCGCATCGTGTGCGCGCTTCTCACCAGCGCGCTTCGAGAC 19212
QY 307 GCGTACACGCCCTACCATCGAGGACTTCCACCGCAAGTTCCTACTCCATCCGCGCGAGGTC 366
|||||
Db 19213 GCGTACACGCCCTACCATCGAGGACTTCCACCGCAAGTTCCTACTCCATCCGCGCGAGGTC 19272
QY 367 TACCAGCTCGACATCTCGACAGCTCCGCGCAACACCGCTTCCCGCGCATCGCGCGCCTC 426
|||||
Db 19273 TACCAGCTCGACATCTCGACAGCTCCGCGCAACACCGCTTCCCGCGCATCGCGCGCCTC 19332
QY 427 TCCATCCTCAC----- 437
|||||
Db 19333 TCCATCCTCACAGGTGAGCGCGGCGCGGAGGTGCGGAGGGAAGCGCGGGAACCTT 19392
QY 438 ----- 437
Db 19393 CGGCCAGGCGCGCGGAGCGCGCTCGGCTGCGCGCGCGGAGTAGTGCCTTCGCG 19452
QY 438 ----- 437
Db 19453 CTTAGAGGCTAGCGCGCGCGCGGCGCTCAAAAGTCAGCGCCGACTTGTCCCTTGGCG 19512
QY 438 -----AGGAGAGCTTTTCATCCT 455
|||||
Db 19513 GCCACCTCACCTTCTCCTTTCTGCTCTCTGTCGCCCTCTAGGAGAGCTTTTCATCCT 19572
QY 456 GGTGTTCACTGTCGACACCGGAGCTCTCTGAGAGGTGACGCGCTCAGCGAGCAT 515
|||||
Db 19573 GGTGTTCACTGTCGACACCGGAGCTCTCTGAGAGGTGACGCGCTCAGCGAGCAT 19632
QY 516 CCTCGACACCAAGTCTTGCTCCTCAAGAACAAACCAAGA--GAAGCTGGAGTGCCTCG 573
|||||
Db 19633 CCTCGACACCAAGTCTTGCTCCTCAAGAACAAACCAAGAANGAGTGGAGTGCCTCG 19692
QY 574 GTCATCTGGGCAACAAGGTTGACCGGACTTCTACCGGAGGTGACCGAGCGGAGATC 633
|||||
Db 19693 GTCATCTGGGCAACAAGGTTGACCGGACTTCTACCGGAGGTGACCGAGCGGAGATC 19752
QY 634 GAGCAGCTGGTGGGCGAGACCCCGAGCGCTGCGCTACTTCGAGATCTCGGCCAAGAAG 693
|||||
Db 19753 GAGCAGCTGGTGGGCGAGACCCCGAGCGCTGCGCTACTTCGAGATCTCGGCCAAGAAG 19812
QY 694 AACAGCAGCTGGACAGATGTTCCGCGCGCTTTCGCGCATGGCCAAAGCTGCCAGCGAG 753
|||||
Db 19813 AACAGCAGCTGGACAGATGTTCCGCGCGCTTTCGCGCATGGCCAAAGCTGCCAGCGAG 19872
QY 754 ATGAGCCGAGACTGCACCGCAAGTCTCGGTGAGTACTCGGAGCTGCTGCACAAGAAG 813
|||||
Db 19873 ATGAGCCGAGACTGCACCGCAAGTCTCGGTGAGTACTCGGAGCTGCTGCACAAGAAG 19932
QY 814 GCGCTGCGGACAAAGACTGCTGGGCGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 873
|||||
Db 19933 GCGCTGCGGACAAAGACTGCTGGGCGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 19992
QY 874 GACGCTTTGGCATGCTGGCACCTTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCATG 933
|||||
Db 19993 GACGCTTTGGCATGCTGGCACCTTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCATG 20052
QY 934 TACATCCGAGAGGCATC 993
|||||
Db 20053 TACATCCGAGAGGCATC 20112
QY 994 AGCTAGGAGCG 1053
|||||
Db 20113 AGCTAGGAGCG 20172
QY 1054 TCCAAAGCGCGCGTGGCGT 1113
|||||

Db 20173 TCCAAAGCGCGCGTGGCGT 20232
QY 1114 CGATCCGTCC 1173
Db 20233 CGATCCGTCC 20292
QY 1174 GGAAGGAAGAGAGCGT 1233
Db 20293 GGAAGGAAGAGAGCGT 20352
QY 1234 CCGTCA 1293
Db 20353 CCGTCA 20412
QY 1294 AGACTTAAGAGTGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGT 1352
Db 20413 AGACTTAAGAGTGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGT 20472
QY 1353 CTTCAAGAAACCGTATCGCGCGTTCGAGGTCAGGACCGCGCGCGCGCGCGCGCGTCTGT 1412
Db 20473 CTTCAAGAAACCGTATCGCGCGTTCGAGGTCAGGACCGCGCGCGCGCGCGCGTCTGT 20532
QY 1413 GATTCCGCGGTTCGTCGACAGCGCGTTCGAGGTCAGGACCGCGCGCGCGCGCGCGT 1472
Db 20533 GATTCCGCGGTTCGTCGACAGCGCGTTCGAGGTCAGGACCGCGCGCGCGCGCGCGT 20592
QY 1473 CGTGGGTCAAAATACATAGCAAGTTCGAGGTCAGGACCGCGCGCGCGCGCGCGT 1532
Db 20593 CGTGGGTCAAAATACATAGCAAGTTCGAGGTCAGGACCGCGCGCGCGCGCGT 20652
QY 1533 ACAAAACAAACCTTGCACCTTTAAGCGTAGTTCGCGTTCGCGTTCGCGTTCGCGTTCG 1592
Db 20653 ACAAAACAAACCTTGCACCTTTAAGCGTAGTTCGCGTTCGCGTTCGCGTTCGCGTTCG 20712
QY 1593 TTACCCAGGTGTTTATCTACTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 1652
Db 20713 TTACCCAGGTGTTTATCTACTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 20772
QY 1653 TTTTAAATATACATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1687
Db 20773 TTTTAAATATACATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 20807

RESULT 8
AC090608
LOCUS AC090608 183598 bp DNA linear HTG 22-MAR-2002
DEFINITION Homo sapiens chromosome 17 clone RP11-524F11 map 17, *** SEQUENCING
IN PROGRESS ***, 2 unordered pieces.
AC090608
VERSION HTG; HTGS_PHASE1; HTGS_FULLTOP; HTGS_ACTIVEFIN.
KEYWORDS Homo sapiens.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 183598)
AUTHORS Birren, B., Linton, L., Nusbaum, C., and Lander, E.
TITLE Homo sapiens chromosome 17, clone RP11-524F11
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 183598)
AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, S.,
Barn, N., Bastien, V., Boguslavsky, L., Boukhalter, B., Brown, A.,
Cammarata, J., Campopiano, A., Choepel, Y., Colangelo, M., Collins, S.,
Collamore, A., Cooke, P., DeArelano, K., Dewar, K., Diaz, J.S.,
Dodge, S., Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J.,
Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-pierre, N.,
Hagos, B., Hearford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,
Jones, C., Karatas, A., LaRoque, K., Lamazares, R., Landers, T.,
Lehoczky, J., Levine, R., Liu, G., MacLean, C., Macdonald, P.,
Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K.,
McPheeters, R., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V.,
Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C.H.,

QY 754 ATGAGCCAGACCTGCACCCGACAGTCTGCTGCTAGTCTGACGCTGCTGACAGAG 813
 DB 601 ATGAGCCAGACCTGCACCCGACAGTCTGCTGCTAGTCTGACGCTGCTGACAGAG 660
 QY 814 GCGCTGGGAAACAAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 873
 DB 661 GCGCTGGGAAACAAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
 QY 874 GAGCCCTTTGGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 933
 DB 721 GAGCCCTTTGGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
 QY 934 TACATCCGCGAGAGAGCGCCGCGGAGCGCCAGCGCCAGAGAGAGAGAGAGAGAG 993
 DB 781 TACATCCGCGAGAGAGCGCCGCGGAGCGCCAGCGCCAGAGAGAGAGAGAGAGAG 840
 QY 994 AGCTAG 1053
 DB 841 AGCTAG 899
 QY 1054 TCCAAAGGAG 1113
 DB 900 TCCAAAGGAG 959
 QY 1114 CGATCCGCGCGAG 1172
 DB 960 CGATCCGCGCGAG 1018
 QY 1173 CGGAG 1232
 DB 1019 CGGAG 1078
 QY 1233 GCGCGCGAG 1292
 DB 1079 GCGCGCGAG 1138
 QY 1293 AAG 1341
 DB 1139 AAG 1187
 RESULT 10
 AF172846 979 bp mRNA linear PRI 05-JAN-2000
 LOCUS AF172846 Homo sapiens ras-related protein (DEXRAS1) mRNA, complete cds.
 DEFINITION AF172846.1 GI:6014488
 VERSION AF172846.1 GI:6014488
 KEYWORDS
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 979)
 AUTHORS Tu, Y. and Wu, C.
 TITLE Cloning, expression and characterization of a novel human
 Ras-related protein that is regulated by glucocorticoid hormone
 JOURNAL Biochim. Biophys. Acta 1489 (2-3), 452-456 (1999)
 MEDLINE 20135605
 PUBMED 10673050
 REFERENCE 2 (bases 1 to 979)
 AUTHORS Tu, Y. and Wu, C.
 TITLE Direct Submision
 JOURNAL Submitted (27-JUL-1999) Department of Cell Biology, University of
 Alabama at Birmingham, 1670 University Blvd., Birmingham, AL 35294,
 USA
 FEATURES
 source location/Qualifiers
 1. 979 /organism="Homo sapiens"
 gene /db_xref="taxon:9606"
 CDS 1. 979 /gene="DEXRAS1"
 64. 909

BASE COUNT 195 a 347 c 284 g 153 t
 ORIGIN
 Query Match 52.28; Score 940.4; DB 9; Length 979;
 Best Local Similarity 99.58; Pred. No. 1.8e-114;
 Matches 975; Conservative 0; Mismatches 1; Indels 4; Gaps 3;
 QY 90 CGAGCCGAG 149
 DB 1 CGAGCCGAG 59
 QY 150 CCCAATGAAGTGGCCGCGAGATGATCAAGAGATGCTGCCAGAGAGAGAGAGAGAG 209
 DB 60 CCCAATGAAGTGGCCGCGAGATGATCAAGAGATGCTGCCAGAGAGAGAGAGAGAG 119
 QY 210 CCGGCGCAAG 269
 DB 120 CCGGCGCAAG 179
 QY 270 CATGCTGCGAG 329
 DB 180 CATGCTGCGAG 239
 QY 330 CTTCCACCGCAAGTCT 389
 DB 240 CTTCCACCGCAAGTCT 299
 QY 390 GTCCGCAACCAACCGCTTCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 449
 DB 300 GTCCGCAACCAACCGCTTCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 359
 QY 450 CATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 509
 DB 360 CATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 419
 QY 510 GCAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 569
 DB 420 GCAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 479
 QY 570 CTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 629
 DB 480 CTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 539
 QY 630 GATGAGACAG 689
 DB 540 GATGAGACAG 599
 QY 690 GAGAGACAG 749
 DB 600 GAGAGACAG 659
 QY 750 GAGAGACAG 809
 DB 660 GAGAGACAG 719
 QY 810 GAGAGACAG 869
 DB 720 GAGAGACAG 779

QY 870 GGGGAGACCTTGGCATGCTGGACCTTGGCGCGCCGCCAGCGTACAGGACCT 929
 Db 780 GGGGAGACCTTGGCATGCTGGACCTTGGCGCGCCGCCAGCGTACAGGACCT 839
 QY 930 CATGTACATCCGGAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 989
 Db 840 CATGTACATCCGGAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 899
 QY 990 CATGACCTGAGAGACCC--GCCGCGCTGGCGA-CACAACTTAAGAGAGACCTTTTGT 1046
 Db 900 CATGACCTGAGAGACCCCGCGCGCTGGCGCGCGCGCGCGCGCGCGCGCGCG 959
 QY 1047 AGTCAATCCAGCGCGCGCG 1066
 Db 960 AGTCAATCCAGCGCGCG 979

RESULT 11

AF498923

LOCUS

DEFINITION Homo sapiens activator of G protein signaling (RASD1) mRNA, PRT 01-MAY-2002

ACCESSION

AF498923

VERSION

AF498923.1

KEYWORDS

GI:20379021

SOURCE

Homo sapiens.

ORGANISM

Homo sapiens

REFERENCE

Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo. Publ. H.L., Ikeda, S.R. and Aronstam, R.S. Submitted (05-APR-2002) CDNA Resource Center, Guthrie Research Institute, One Guthrie Square, Sayre, PA 18840, USA

FEATURES

source

location/Qualifiers

1..846

/organism="Homo sapiens"

/db_xref="taxon:9606"

/tissue_type="brain"

1..846

/gene="RASD1"

/note="RASD1"

1..846

/gene="RASD1"

/codon_start=1

/product="activator of G protein signaling"

/protein_id="AA021071.1"

/db_xref="GI:20379021"

/translation="MKLAKMKKPCSPDSELSIPKNCRYMVLIGSSKVGKTAIVSR LIGREDATPTIEDHFRKFSIRGEVQLDITSGNHPFAMRRLSILGDFVITY FSLNDSFEVEQRLROOILIDYKSLKTKENVDPVLYICNGKDRDFEVQOREI EOLVDDPQRCAYFEISAKKNSLDPMRALFAMAKLPSEMSPLDHRKVSVOYDVLIH KRALNKKLLRAGSGGGGDPDAGIYAPFARRPSVSDIMYIRKASAGSOAKDE RCVIS"

BASE COUNT 176 a 288 c 252 g 130 t
 ORIGIN

Query Match 47.0%; Score 846; DB 9; Length 846;
 Best Local Similarity 100.0%; Pred. No. 4, 6e-102; Indels 0; Gaps 0;

Matches 846; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 154 ATGAACCTGGCGCGCATGATCAAGAGATGTCGCGAGGACTCGAGCTGATCCCG 213
 Db 1 ATGAACCTGGCGCGCATGATCAAGAGATGTCGCGAGGACTCGAGCTGATCCCG 60
 QY 214 GCCAAGACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 273
 Db 61 GCCAAGACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120
 QY 274 GTGTGGCGCTTCTGACCGCGCGCTTGAAGAGAGCGCTACAGGCTACCATGAGGACTTC 333
 Db 121 GTGTGGCGCTTCTGACCGCGCGCTTGAAGAGAGCGCTACAGGCTACCATGAGGACTTC 180

QY 334 CACCGAAGTTTACTTCATCCGCGCGCGAGGTCTACAGCTGACATCTCGACAGCTC 393
 Db 181 CACCGAAGTTTACTTCATCCGCGCGCGAGGTCTACAGCTGACATCTCGACAGCTC 240
 QY 394 GGCACACACCGTTTCCCGCGCATGGGCGCGCTCTCTCATCTCTACAGGAGCTTTTATC 453
 Db 241 GGCACACACCGTTTCCCGCGCATGGGCGCGCTCTCTCATCTCTACAGGAGCTTTTATC 300
 QY 454 CTGGTTTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 513
 Db 301 CTGGTTTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360
 QY 514 ATCTTGACACCAAGTCTTCTCTCAAGAACCAAGAACCAAGAACCAAGAACCAAG 573
 Db 361 ATCTTGACACCAAGTCTTCTCTCAAGAACCAAGAACCAAGAACCAAGAACCAAG 420
 QY 574 GTTCATCTGGCGCAAGAGGTGACCGCGCGCTTCAACCGCGCGAGTACAGCGAGATC 633
 Db 421 GTTCATCTGGCGCAAGAGGTGACCGCGCGCTTCAACCGCGAGTACAGCGAGATC 480
 QY 634 GAGCAGCTGTGTGGCGAGCAGACCCCGCGCGCTTCAACCGCGAGTACAGCGAGATC 693
 Db 481 GAGCAGCTGTGTGGCGAGCAGACCCCGCGCGCTTCAACCGCGAGTACAGCGAGATC 540
 QY 694 AACACAGAGCTGTGACAGATGTTCCGCGCGCTTCTTCCGATGCGCAAGCTGCGAGAG 753
 Db 541 AACACAGAGCTGTGACAGATGTTCCGCGCGCTTCTTCCGATGCGCAAGCTGCGAGAG 600
 QY 754 ATGACCCAGACCTGACACCCAGAGTCTGCTGAGTACAGTACAGTACAGTACAGT 813
 Db 601 ATGACCCAGACCTGACACCCAGAGTCTGCTGAGTACAGTACAGTACAGTACAGT 660
 QY 814 GCGCGCGCAAGAGAGTCTGCTGAGTACAGTACAGTACAGTACAGTACAGTACAGT 873
 Db 661 GCGCGCGCAAGAGAGTCTGCTGAGTACAGTACAGTACAGTACAGTACAGTACAGT 720
 QY 874 GAGCGCTTGGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 933
 Db 721 GAGCGCTTGGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
 QY 934 TACATCCGCGAAGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 993
 Db 781 TACATCCGCGAAGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 840
 QY 994 AGCTAG 999
 Db 841 AGCTAG 846

RESULT 12

BC034166

LOCUS

DEFINITION Mus musculus, RAS, dexamethasone-induced 1, clone MGC:36188

ACCESSION

BC034166

VERSION

BC034166.1

KEYWORDS

house mouse.

SOURCE

house mouse.

ORGANISM

Mus musculus

REFERENCE

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus. Strausberg, R. Submitted (02-JUL-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

AUTHORS

TITLE

JOURNAL

REMARK

COMMENT

NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
 Contact: MGC help desk
 Email: gcgaps-remail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 CDNA Library Preparation: Life Technologies, Inc.

QY 1639 TCGTTTATTTGGTTTAAATACATAATAATTTAAATGCAAAAAA 1698
Db 1552 -----TTTTTTTAAATAATAATTTAGAAAGGATTAATAAAAAA 1598
QY 1699 AAAAAAAAAAAAAA 1712
Db 1599 AAAAAAAAAAAAAA 1612

RESULT 13
AF009246 1623 bp mRNA linear ROD 13-FEB-1998
LOCUS AF009246
DEFINITION Mus musculus ras-related protein (DEXRAS1) mRNA, complete cds.
ACCESSION AF009246
VERSION AF009246.1 GI:2253712
KEYWORDS
SOURCE Mus musculus.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE
AUTHORS 1 (bases 1 to 1623)
TITLE Dexamethasone rapidly induces a novel ras superfamily
JOURNAL member-related gene in A3T-20 cells
MEDLINE J. Biol. Chem. 273 (6), 3129-3131 (1998)
PUBMED 9452419
REFERENCE 2 (bases 1 to 1623)
AUTHORS Kempainen, R.J. and Behrend, E.N.
TITLE Direct Submission
JOURNAL Submitted (18-JUN-1997) Physiology & Pharmacology, Auburn
University College of Veterinary Medicine, 213 Greene Hall, Auburn,
AL 36849, USA

FEATURES
source location/Qualifiers
1..1623
/organism="Mus musculus"
/db_xref="taxon:10090"
/cell_line="corticotrope tumor cell line A3T-20"
1..1623
/gene="DEXRAS1"
142..984
/gene="DEXRAS1"
/note="Induced by dexamethasone"
/product="ras-related protein"
/protein_id="AAC5338.1"
/db_xref="GI:2253713"
/translation="KRLAMTKMCPDSELSIPAKNCYRVILIGSSKVKTAIVSR
LTGRDVAITPTIEDFHRKYSIRGEVYQDILDTSGNHPAPARRLSILGVFLLV
FSLNDRDSFEVQRLKQDILDTKSLKTKKTEVDPLVIGNGDSDPYREYQREI
EQLVDDPQRCAYFEISAKNSISLQMRALFAAKLPSEMPDLHRKVSVOYCDVH
KRALNRKILRAGSGGGDHGDAFGLIAPFARRRPSVHSDLMITIKTSVGSQAKDKR
CVIS"

BASE COUNT 412 a 437 c 417 g 357 t
ORIGIN

Query Match 42.4%; Score 764.2; DB 10; Length 1623;
Best Local Similarity 71.9%; Pred. No. 2.7e-91;
Matches 1215; Conservative 0; Mismatches 363; Indels 113; Gaps 12;

QY 33 GAGCCGCGCCAGCCGACAGCCGCTCCAGCCGCTACCCCGCTGCAACCCAGCA 92
Db 24 GAGCCGCGCCAGCCGACAGCCGCTCCAGCCGCTTTC---CGTCACTCCGAGTTC 80
QY 93 CCCTGACCGCTCTCTGCGCTCTCTCGGCGCCGCGCCGCTGCGCCCTGCGC 152
Db 81 TCTGCGCGCTCTCTGCGCTCTCTCTGCGCGCGATCCGCGCCGCGCCCTCTACCC 140
QY 153 AATGAACTGGCGCGGATGATCAAGAGATGCGCGGAGCACTGGAAGTATATCC 212
Db 141 AATGAACTGGCGCGGATGATCAAGAGATGCGCGGAGCACTGGAAGTATATCC 200
QY 213 GCCCAAGAACTGCTATGCGATGATCATCTCGGCTCGTCCAGGTGGGCAAGCGGCAT 272

Db 201 GCCCAAGAACTGCTATGAGATGATGATCTCTGCGCTATCCAAATGGGCAAGCGGCAT 260
QY 273 CGTGTGCGCTTCTCTACCGCGCGCTTCGAGAGCGCTTACAGCGCTACATCAGAGACT 332
Db 261 TGTGTGCGCTTCTCTACCGCGCGCTTCGAGAGCGCTTACAGCGCTACATCAGAGACT 320
QY 333 CCACCGCAAGTTTACATCCATCCCGGCGGAGGTATACAGCTGACATCTTCGACAGCTC 392
Db 321 CCACCGCAAGTTTACATCCCGGCGGAGGTATACAGCTGACATCTTCGACAGCTC 380
QY 393 CGGCAACACCGCTTCCCGCATATGCGCGCGCTTCATCTCCTACAGAGAGCTTTTCA 452
Db 381 CGGCAATATCTGTTCCCGCATATGCGCGCGCTTCATCTCCTACAGAGAGCTTTTCA 440
QY 453 CCGTGTTCAGTCTGAGCAACCGGACTCTCTTCGAGAGGTGACGCGCTCAGGAGCA 512
Db 441 TGTGTGTTTCAAGTTAGCAACCGGACTCTTCGAGAGGTGACGCGCTCAGGAGCA 500
QY 513 GATCTCGACACCAAGTCTTGTCTCAGAAACAAACCAAGAGAGAGTGTGAGTGTCCCT 572
Db 501 GATCTCGACACCAAGTCTTGTCTCAGAAACAAACCAAGAGAGAGTGTGAGTGTCCCT 560
QY 573 GGTATCTGCGGCAAGAGGTGACCGGACTTCTACCGGAGGTGAGCAGCGGAGAT 632
Db 561 GGTATCTGCGGTAACAAAGGAGGACCGGACTTCTACCGGAGGTGAGCAGCGGAGAT 620
QY 633 CGAGCAGCTGTGGGCGACAGACCCCGGCTGCGCTTCTACATCTGAGATCTGCGCAAGAA 692
Db 621 TGACACACTGTGTGAGACAGACCCCTCAGCGCTTGTGCTTCTGAGATCTGAGCAAGAA 680
QY 693 GAACAGCAGCTGTGAGACAGATGTTCCGCGCGCTTGTGCGCATGCGCAACCTGCCAGGCA 752
Db 681 GAACAGCAGCTGTGAGACAGATGTTCCGCGCGCTTGTGCGCATGCGCAACCTGCCAGGCA 740
QY 753 GATAGCCCAAGACTGTGACCGCAAGGTCTTCGCTGAGTACTGAGAGTCTCACAAGAA 812
Db 741 GATAGCCCAAGACTGTGACCGCAAGGTATCTGCTGAGTACTGAGAGTCTCACAAGAA 800
QY 813 GGCCTCGGCAAGAAAGTGTGTGCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 872
Db 801 GGCCTCGGCAAGAAAGTGTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 857
QY 873 CGAGCCCTTGGCATGTGTGACCTTTCGCGCGCGGCGGCGGCGGCGGCGGCGGCGG 932
Db 858 CGATGCTTGTGATCTTGGCGCGCTTTCGCGCGGCGGCGGCGGCGGCGGCGGCGG 917
QY 933 GTACATCCGGAAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 992
Db 918 GTACATTCGGAAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 977
QY 993 CAGCTAGAGCCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1052
Db 978 CAGTATGAGAGCCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1037
QY 1053 ATCCAAAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1112
Db 1038 ATGCGATCTCGAGTATGAGTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1094
QY 1113 GCGATCCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1172
Db 1095 -----CTTCCCTGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1149
QY 1173 CGGAGAAAGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1232
Db 1150 TGAAGAGAAAGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1184
QY 1233 GCCCGCAACCCCATTAATCTTGGAGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1292
Db 1185 -----CAGCTCGATCTGAGATGCGGAGAGCGGCGGCGGCGGCGGCGGCGGCGG 1237
QY 1293 AAGACCTAAGAGTGAAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1352

PF 07-MAY-1999; 99WO-US1015


```

Db 1621 TAGGCTTTAAAGTATGCTTTATTTGTTTAAATATCAATTAATTTAA 1680
QY 1681 TGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1740
Db 1681 TGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1740
QY 1741 AAAAAAAAAAGCGCCCTGAGATCATCTAGAGCGCCGATCATGTAATTAGTTATGAA 1800
Db 1741 AAAAAAAAAAGCGCCCTGAGATCATCTAGAGCGCCGATCATGTAATTAGTTATGAA 1800
QY 1801 C 1801
Db 1801 C 1801

```

RESULT 2
AAZ36914
ID AAZ36914 standard; DNA; 1740 BP.

AC AAZ36914;

DE 13-MAR-2000 (first entry)

DE DNA encoding an activator of G protein signalling (AGS) protein.

KW Activator of G protein signalling; AGS; ras-related G protein;
KW GTP hydrolysis; G protein activity; pheromone response pathway;
KW G protein-coupled signal transduction; G-gamma selectivity;
KW cellular signal transduction; ss.

OS Homo sapiens.

FT Key Location/Qualifiers
FT CDS 145..991
FT /*tag= a

FT /product= "activator of G protein signalling (AGS)
FT protein"

XX WO958670-A1.

XX 18-NOV-1999.

XX 07-MAY-1999; 99WO-0510151.

XX 08-MAY-1998; 98US-0084842.

XX 07-OCT-1998; 98US-0103355.

XX (CADU-) CADUS PHARM CORP.

XX Cismowski M, Duzic E;

XX WPI: 2000-072337/06.

XX P-PsDB; AAY33924.

PT A new activator of G protein signalling used to treat disorders
PT characterized by an aberrant AGS protein activity -

PS Disclosure: Page 146-148; 162pp: English.

CC The present sequence encodes an activator of G protein signalling (AGS)
CC protein. The cDNA sequence was isolated from a human liver cDNA
CC library. The AGS protein exhibits homology to ras-related G proteins,
CC and contains alterations in conserved amino acids consistent with a
CC deficiency in GTP hydrolysis activity. AGS stimulates G protein
CC activity. G protein-coupled signal transduction and the pheromone
CC response pathway in a receptor-independent manner. The AGS protein
CC also shows G-gamma selectivity, as measured by growth assays in
CC yeast expressing various mammalian G-gamma constructs, and
CC tissue-specific expression, as measured by Northern blot analysis.
CC The AGS protein can be used to screen for compounds that modulate
CC cellular signal transduction. The protein is used to treat disorders
CC characterized by an aberrant AGS protein activity or AGS nucleic acid

```

CC expression.
XX
SO Sequence 1740 BP; 422 A; 546 C; 483 G; 289 T; 0 other;
Query Match 96.8%; Score 1740; DB 21; Length 1740;
Best Local Similarity 100.0%; Pred. No. 6.2e-230;
Matches 1740; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 9 GAGGGAGCCGAGCCCAAGCCGACCCGCGCCAGAGAGAGCCCTCCACCGG 68
Db 1 GAGGGAGCCGAGCCCAAGCCGACCCGCGCCAGAGAGAGAGCCCTCCACCGG 60
QY 69 TCACCCCGCGTGCACACCCGACCCCTGACCGGCTGTGCCCTTCTCGGCCCGG 128
Db 61 TCACCCCGCGTGCACACCCGACCCCTGACCGGCTGTGCCCTTCTCGGCCCGG 120
QY 129 CCGGCCCTCGCGGCCCTCTGCCCCAATGAACTGGCCGCGATGATCAAGAAATGTGCC 188
Db 121 CCGGCCCTCGCGGCCCTCTGCCCCAATGAACTGGCCGCGATGATCAAGAAATGTGCC 180
QY 189 GAGGAGCTGGAGCTGAGTATCCCGGCAAGAACTGTATCGATGTATCTCGGCTC 248
Db 181 GAGGAGCTGGAGCTGAGTATCCCGGCAAGAACTGTATCGATGTATCTCGGCTC 240
QY 249 GTCCAAAGTGGGCAAGAGCGCATCGTGTGCGCTTCTACCGCGCGTTCGAGAGCG 308
Db 241 GTCCAAAGTGGGCAAGAGCGCATCGTGTGCGCTTCTACCGCGCGTTCGAGAGCG 300
QY 309 CTACAGGCTTACCATGAGAGACTTCCACCGCAAGTTCTACTCATCCGCGGAGGTCTA 368
Db 301 CTACAGGCTTACCATGAGAGACTTCCACCGCAAGTTCTACTCATCCGCGGAGGTCTA 360
QY 369 CCAGCTCGACATCTCTGACACGTCGCGCAACCGGCTTCCCGCATCGCGGCTCTC 428
Db 361 CCAGCTCGACATCTCTGACACGTCGCGCAACCGGCTTCCCGCATCGCGGCTCTC 420
QY 429 CATCTTCACAGAGAGACGTTTCTATCTGTGTGTCAGTGTGACAAACCGGACTCTTGA 488
Db 421 CATCTTCACAGAGAGACGTTTCTATCTGTGTGTCAGTGTGACAAACCGGACTCTTGA 480
QY 489 GGAGGTGCAAGCGGCTCAGGACAGATCTCTGACACCAAGTCTTGGCTTAACAAAC 548
Db 481 GGAGGTGCAAGCGGCTCAGGACAGATCTCTGACACCAAGTCTTGGCTTAACAAAC 540
QY 549 CAAGGAGAACGTGAGACGTCCTCTGTCATCTGCGGCAACAGGAGGACCGGACTCTGA 608
Db 541 CAAGGAGAACGTGAGACGTCCTCTGTCATCTGCGGCAACAGGAGGACCGGACTCTGA 600
QY 609 CCGCGAGGTGACACAGCGCGAGATCGAGACGCTGTGGGCGACGACCCCAAGCGTGGCG 668
Db 601 CCGCGAGGTGACACAGCGCGAGATCGAGACGCTGTGGGCGACGACCCCAAGCGTGGCG 660
QY 669 CTACTTCGAGATCTCGGCCAAGAGACAGCCCTGACACAGATGTTCCGCGGCTTT 728
Db 661 CTACTTCGAGATCTCGGCCAAGAGACAGCCCTGACACAGATGTTCCGCGGCTTT 720
QY 729 CCGCATGGCCAAAGCTGCCAGGAGATGAGCCACAGCTGACCGCAAGGTTCTGGTGA 788
Db 721 CCGCATGGCCAAAGCTGCCAGGAGATGAGCCACAGCTGACCGCAAGGTTCTGGTGA 780
QY 789 GTACTGCGACGCTGTGACACAAAGAGCGCTGTGGGAAAGAAAGTCTGTGGGCGGAG 848
Db 781 GTACTGCGACGCTGTGACACAAAGAGCGCTGTGGGAAAGAAAGTCTGTGGGCGGAG 840
QY 849 CCGCGCGCGCGCGCGAGACCCGCGAGCGCTTGGATCGATGAGACCTTCTCGGCGCGG 908
Db 841 CCGCGCGCGCGCGCGAGACCCGCGAGCGCTTGGATCGATGAGACCTTCTCGGCGCGG 900
QY 909 GCCCAGCTTACACAGCACTTCATGTATCATCCGCGAGAAAGCCAGCGCGGACGACGCG 968
Db 901 GCCCAGCTTACACAGCACTTCATGTATCATCCGCGAGAAAGCCAGCGCGGACGACGCG 960
QY 969 CAAGGACAAAGAGCGCTGTGATCATGATAGAGCGCCGCGGCTGCGACACAACTTA 1028

```



```

QY 427 TCCATCTCAGAGAGACGTTTTCATCTCTGTGTTCACTGTGAGACACCGGACGCTCTTC 486
DB 453 TCCATCTCAGAGAGACGTTTTCATCTCTGTGTTCACTGTGAGACACCGGACGCTCTTC 512
QY 487 GAGAGAGTGCACGCGCTCAGGAGACGATCCGACACACACCTCTTGGCTTCAAGAACAA 546
DB 513 GAGAGAGTGCACGCGCTCAGGAGACGATCCGACACACACCTCTTGGCTTCAAGAACAA 572
QY 547 ACCAAGAGAGACGTTGAGCGCTGCTGTATCTGCGGACACAGAGGATGACCGGACCTTC 606
DB 573 ACCAAGAGAGACGTTGAGCGCTGCTGTATCTGCGGACACAGAGGATGACCGGACCTTC 632
QY 607 TACCGGAGGTGAGACGCGGACGAGATCGACAGCTGTGTGGCGAGACGCCCGGCGTGC 666
DB 633 TACCGGAGGTGAGACGCGGACGAGATCGACAGCTGTGTGGCGAGACGCCCGGCGTGC 692
QY 667 GCCTACTTCGAGATCTCGGCGCAGAGAACAGACGCTGTGACCGAGATGTCGGCGGCTC 726
DB 693 GCCTACTTCGAGATCTCGGCGCAGAGAACAGACGCTGTGACCGAGATGTCGGCGGCTC 752
QY 727 TTCGCGATGCGCCAGCTGCCACGAGATGAGCCGACCTGCAACCGCAAGGTCTCGGTG 786
DB 753 TTCGCGATGCGCCAGCTGCCACGAGATGAGCCGACCTGCAACCGCAAGGTCTCGGTG 812
QY 787 CAGTACTGCGAGCTGCTGACAGAGAGGCGTGGGGAACAGAGCTGTCTCGGCGGCG 846
DB 813 CAGTACTGCGAGCTGTCTGACAGAGAGGCGTGGGGAACAGAGCTGTCTCGGCGGCG 872
QY 847 AGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 906
DB 873 AGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 932
QY 907 CGGCGCGAGCTGACAGAGAGCTATGATCTCGCGGAGAGGCGGCGGCGGCGGCGGCGG 966
DB 933 CGGCGCGAGCTGACAGAGAGCTATGATCTCGCGGAGAGGCGGCGGCGGCGGCGGCGG 992
QY 967 GCCAAGAGAGAGAGAGCTGCTGATCAGTACAGAGAGGCGGCGGCGGCGGCGGCGGCG 1026
DB 993 GCCAAGAGAGAGAGAGCTGCTGATCAGTACAGAGAGGCGGCGGCGGCGGCGGCGGCG 1052
QY 1027 TAAAGAGAGACCTTTTGTAAATCAAAATCCAGCGCGGCGGCGGCGGCGGCGGCGG 1086
DB 1053 TAAAGAGAGACCTTTTGTAAATCAAAATCCAGCGCGGCGGCGGCGGCGGCGGCGGCG 1112
QY 1087 CGGCGGAGTGGGCGTCCCTCCCGCGGAGATCCGCGCGGCGGCGGCGGCGGCGGCGGCG 1146
DB 1113 CGGCGGAGTGGGCGTCCCTCCCGCGGAGATCCGCGCGGCGGCGGCGGCGGCGGCGGCG 1172
QY 1147 AACCGAGAGAGAGAGCTGCTGCTCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1206
DB 1173 AACCGAGAGAGAGAGCTGCTGCTCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1232
QY 1207 TTTCCCGACCCCGGCTCCCGATGAGGCGCGGCAACCCCACTTTTGGAGGCGAGGCGC 1266
DB 1233 TTTCCCGACCCCGGCTCCCGATGAGGCGCGGCAACCCCACTTTTGGAGGCGAGGCGC 1292
QY 1267 CAGCGGAGGTGATTTATCTTCTCAAGAGCTAAGAGTATGAGGCGGCGGCGGAGAGGA 1326
DB 1293 CAGCGGAGGTGATTTATCTTCTCAAGAGCTAAGAGTATGAGGCGGCGGCGGAGAGGA 1352
QY 1327 TGTGAAGTATTCAGAGCTGCTGAGGCTTCAAGAAACCGCTGATCGCGGCTTGAAGGCTAG 1386
DB 1353 TGTGAAGTATTCAGAGCTGCTGAGGCTTCAAGAAACCGCTGATCGCGGCTTGAAGGCTAG 1412
QY 1387 GACCCAGGGGCGATTAATCTGTCTGTGATTCGCGGTTGCTGTGACAGCGGTTAGAGCCCTC 1446
DB 1413 GACCCAGGGGCGATTAATCTGTGTGATTCGCGGTTGCTGTGACAGCGGTTAGAGCCCTC 1472
QY 1447 TGGCCTCCCAAACTAAGCGGGGCGGCTGTGCAATATACCAAGAGACTTGTATTAC 1506
DB 1473 TGGCCTCCCAAACTAAGCGGGGCGGCTGTGCAATATACCAAGAGACTTGTATTAC 1532

```

```

QY 1507 ATGTGAGTGAAGTGCACAAAGAACACAAACAACTTGCACTTTAAGGTAGTTCCG 1566
DB 1533 ATGTGAGTGAAGTGCACAAAGAACACAAACAACTTGCACTTTAAGGTAGTTCCG 1592
QY 1567 GTGTCAACATGAGACGAGCAACAAACCTTACCGAGTGTATTAAGTGTGTGAGGT 1626
DB 1593 GTGTCAACATGAGACGAGCAACAAACCTTACCGAGTGTATTAAGTGTGTGAGGT 1652
QY 1627 CTTTAAAGTATTTCTTATTTTGTGTTTATTAATTAATTAATTAATTAATTAAT 1686
DB 1653 CTTTAAAGTATTTCTTATTTTGTGTTTATTAATTAATTAATTAATTAATTAAT 1712
QY 1687 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1746
DB 1713 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1772
QY 1747 AA 1748
DB 1773 AA 1774

```

RESULT 4

AAZ23022

AAZ23022 standard; cDNA; 1841 BP.

AAZ23022;

17-JAN-2000 (first entry)

Human kd312 polypeptide encoding cDNA.

kd312 polypeptide; cancer; immunodeficiency disorder; AIDS; stroke;

heart attack; head trauma; neurodegenerative disease; human;

Parkinson's disease; Alzheimer's disease; ss.

Homo sapiens.

W0950288-A2.

07-OCT-1999.

30-MAR-1999; 99WO-US06993.

31-MAR-1999; 98US-0053374.

(AMGE-) AMGEN INC.

Yen K;

WPI: 1999-601322/51.

P-PSDB: AAY42693.

kd312 polypeptides useful for treating diseases and disorders

associated with alterations in cell proliferation and cell death

Claim 1; Fig 8; 85pp; English.

PS

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

```

108 CGAGCGAAGCGGAGCGCCAGAGCCGCGCCAGCCGAGAGAGCGCTCCAGCC 167
109 GCTACACCGCGCGTCCAGCCAGCGACCTCAGCGCTCTCTGCGCTTCTCTGCGCCG 126
110 GCTACACCGCGCGTCCAGCCAGCGACCTCAGCGCTCTCTGCGCTTCTCTGCGCCG 126
168 GCTACACCGCGCGTCCAGCCAGCGACCTCAGCGCTCTCTGCGCTTCTCTGCGCCG 227
127 CGCGCGCGCGTCCAGCGCGCTTCCAGCGCGTCCAGCGCGTCCAGCGCGTCCAGCGCG 186
128 CGCGCGCGCGTCCAGCGCGCTTCCAGCGCGTCCAGCGCGTCCAGCGCGTCCAGCGCG 287
187 CGAGGAGCTCGAGAGCTGATATCCCGCGCAAGAACTGCTATCGATGCTATCTCGCGC 246
288 CGAGGAGCTCGAGAGCTGATATCCCGCGCAAGAACTGCTATCGATGCTATCTCGCGC 347
247 TCGTCCAGAGTGGGCAAGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 306
348 TCGTCCAGAGTGGGCAAGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 407
307 GCGTACACGCTACGAGCTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 366
408 GCGTACACGCTACGAGCTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 467
468 TACGAGTGGAGTCCGCGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 527
427 TCGATCTCTACAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 486
528 TCGATCTCTACAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 587
487 GAGAGAGTGGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 546
588 GAGAGAGTGGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 647
547 ACCAAGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 606
648 ACCAAGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 707
607 TACCGGAGAGTGGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 666
708 TACCGGAGAGTGGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 767
667 GCGTACTCTGAGATCTGCGCGCAAGAAAGAGAGCTGCTGCTGCTGCTGCTGCTGCTG 726
768 GCGTACTCTGAGATCTGCGCGCAAGAAAGAGAGCTGCTGCTGCTGCTGCTGCTGCTG 827
727 TTTGCGCATGGCGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 786
828 TTTGCGCATGGCGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 887
787 CAGTACGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 846
888 CAGTACGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 947
847 AGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 906
948 AGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1007
907 CGCGCGCGCGTACACAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 966
1008 CGCGCGCGCGTACACAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1067
967 GCGAAGAGAGAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1026
1068 GCGAAGAGAGAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1127
1027 TAAGGAGAGAGCTTTTGTAAATCAATCAAGCGCGCGCGCGCGCGCGCGCGCGCGCG 1086
1128 TAAGGAGAGAGCTTTTGTAAATCAATCAAGCGCGCGCGCGCGCGCGCGCGCGCGCG 1187
1087 CGCGCGAGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1146

```

```

1188 CGCGCGAGCTGCGCTCTCCCTCCCGCGATCCGCGCGCGCGCGCGCGCGCGCGCGCTG 1247
1147 AACGAGAGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1206
1248 AACGAGAGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1307
1207 TTTCCGACCGCGCGTCCCGCATGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1266
1308 TTTCCGACCGCGCGTCCCGCATGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1366
1267 CAGCGAGAGAGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1326
1367 CAGCGAGAGAGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1426
1327 TGTGAAGTATTCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1386
1427 TGTGAAGTATTCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1486
1387 GACCCAGCGGCGATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1446
1487 GACCCAGCGGCGATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1546
1447 TCCCTCTCCGGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1506
1547 TCCCTCTCCGGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1606
1507 ATGTAGTGAAGTATTCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1566
1607 ATGTAGTGAAGTATTCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1666
1567 GTGTCAACATGACACGAGCAAAACCTTACCAGAGTGTATACGTGTGTGTGTGTGTGT 1626
1667 GTGTCAACATGACACGAGCAAAACCTTACCAGAGTGTATACGTGTGTGTGTGTGTGT 1726
1627 CTTTAAAGTATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTATGCTT 1686
1727 CTTTAAAGTATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTATGCTT 1786
1687 AAAAAAAAAA 1698
1787 AAAAAAAAAA 1798

RESULT 5
AAZ23023
ID AAZ23023 standard; DNA; 3986 BP.
XX
AC AAZ23023;
XX
DT 17-JAN-2000 (first entry)
XX
DE Human kd312 genomic DNA sequence.
XX
KW kd312 polypeptide; cancer; immunodeficiency disorder; AIDS; stroke;
heart attack; head trauma; neurodegenerative disease; human;
Parkinson's disease; Alzheimer's disease; ss.
XX
OS Homo sapiens.
XX
PN WO950288-A2.
XX
PD 07-OCT-1999.
XX
PF 30-MAR-1999; 99WO-US06993.
XX
PR 31-MAR-1998; 98US-0053374.
XX
PA (AMGE-) AMGEN INC.
XX
PI Yen K;
XX
DR WPI; 1999-601322/51.
DR P-PSDB; AAY42693.

```


Db 1085 TCGCGCAAGTGGCCACCAAGGCGGAGCTGCTGGTGTGCGGCGA---CGAAGACTGC 1141
 QY 667 GCCTACTTCAGAGATCGCGCCAGAGAGACAGAGCCTTGAGACCAAGATGTTCCGCGCGTC 726
 Db 1142 GCTTACTTCGAGGTGTCCGCGCAAGAGAACACAAAGTGGAGATGTTCTAGTGTCTC 1201
 QY 727 TTGCGCATGGCCAGCTCCAGGAGATGAGCCCAAGCTCGACCCGAGAGTCTCGGTG 786
 Db 1202 TTCAAGATGGCCAGCTCCAGAGAGATGAGCCCGCCCTGCAATCCAGATCTCGGTG 1261
 QY 787 CAGTACTGCGACGTGTCGACAGAGAGCGCTGCGGAAACAAGATGCTCGGCGCGCG 846
 Db 1262 CAGTACGGGTGACGCTTCCACCCCAAGGCGC-----C 1291
 QY 847 AGCGGCGCGCGCGCGCGAGACCCGCGGAGCGCTTTGGCATCTGCGACCCCTTCGCGCG 906
 Db 1292 TTTCGATGCGCGCGCGCGAGAGATGAGCGCCCTATGCGATGCTGCGCGCGCGCG 1351
 QY 907 CGCGCCAGCGCTACACAGCAGCTTATCTACATCCGCGAGAGAGCGCGCGCGCGCGAG 966
 Db 1352 CGCGCCAGCGCTACACAGCTTATCTACATCCAGGCGCAAGGTCTTGGGAGAGCGCGAG 1411
 QY 967 GCCAAGAGCAAGAGAGCGCTGCTGATC 993
 Db 1412 GCCGTGAGAGGAGCAAGTGCACCATC 1438

RESULT 11

AA236913

ID AA236913 standard; DNA: 837 BP.

XX AA236913:

DT 13-MAR-2000 (first entry)

DE DNA encoding a homologue of activator of G protein signalling AGS1.

KW Activator of G protein signalling; AGS; ras-related G protein; GTP hydrolysis; G protein activity; pheromone response pathway;

KM G protein-coupled signal transduction; G-gamma selectivity;

KW cellular signal transduction; AGS1 homologue; ss.

XX Homo sapiens.

FH Key Location/Qualifiers

FT CDS 1..837

FT /tag= a

FT /product= "AGS1 homologue"

XX MO958670-11.

XX 18-NOV-1999.

XX 07-MAY-1999; 99MO-US10151.

XX 08-MAY-1998; 98US-0084842.

XX 07-OCT-1998; 98US-0103355.

XX (CADU-) CADUS PHARM CORP.

XX Cismowski M, Duzic E;

XX MPI: 2000-072337/06.

XX P-PSDB: AAY53923.

XX A new activator of G protein signalling used to treat disorders

XX characterized by an aberrant AGS protein activity -

XX Example 15; Page 143-144; 162pp; English.

XX The present sequence encodes a homologue of a human AGS1 (activator

XX of G protein signalling (AGS)) protein. The AGS cDNA sequence was

XX isolated from a human liver cDNA library. The AGS protein exhibits

CC homology to ras-related G proteins, and contains alterations in
 CC conserved amino acids consistent with a deficiency in GTP hydrolysis
 CC activity. AGS stimulates G protein activity, G protein-coupled signal
 CC transduction and the pheromone response pathway in a receptor-independent
 CC manner. The AGS protein also shows G-gamma selectivity, as measured by
 CC growth assays in yeast expressing various mammalian G-gamma constructs,
 CC and tissue-specific expression, as measured by Northern blot analysis.
 CC The AGS protein can be used to screen for compounds that modulate
 CC cellular signal transduction. The protein is used to treat disorders
 CC characterized by an aberrant AGS protein activity or AGS nucleic acid
 CC expression.

Sequence 837 BP; 177 A; 272 C; 233 G; 155 T; 0 other;

Query Match 22.6%; Score 407.4; DB 21; Length 837;

Best Local Similarity 71.5%; Pred. No. 1.8e-47;

Matches 594; Conservative 0; Mismatches 201; Indels 36; Gaps 3;

QY 166 GCGATGATCAAGATGTGCGCGAGCGAGCTGAGTATCCCGGCAAGAACTGC 225
 Db 34 GCCATGATGAAGCTTTGTCAGCGGAGACTGACGCTGATGTCGCCCAAAACTCA 93
 QY 226 TATCGATGATCAATCTCTGCGCTGCTCAAGTGGGCAAGCGCATCTGTCGCGCTTC 285
 Db 94 TACCGATGATGATGCTGCTGCTGCTCTCTCGGTGGGCAAGAGCTCATGCTGCTGCTTC 153
 QY 286 CTCACCGCGCGCTTGAGAGAGCGCTTACAGCGCTTACAGCTGAGACTTCCAGCGCAAGTTC 345
 Db 154 CTCATGATGCGCGCTTTGAGAGCAATGACACCGCCATGAGAGCTTCCACCGTAAGCTA 213
 QY 346 TACTCATTCGCGCGAGGTCTACAGCTGACATCTGACACGCTCCGCGCAACACCGC 405
 Db 214 TACACATTCGCGCGAGCATGTACAGCTGACATCTGAGATCTGAGACACACCGC 273
 QY 406 TTTCCCGCGCGCGCGCTCTCCATCTCCAGAGAGAGCTTTTATCTGCTGCTGCTGCT 465
 Db 274 TTTCCCGCGCGCGAGGTGCTTCTCATCTCCAGAGAGAGTCTTCTATCTGCTGCTGCTGCT 333
 QY 466 CTGGAACACCGCGAGCTCTTACAGAGGTGACGCGCTACAGAGACAGATCTCTGAGACGC 525
 Db 334 CTGATTAACCGGAGATCTCTGATGAGTCAAGCGCTTCAAGAGACAGATCTCTGAGGCTC 393
 QY 526 AAGTCTGCTCTAAGAAACAAACAGAGAGAGTGAAGTGGCCCTGATCTGCGGC 585
 Db 394 AAGTCTGCTCTAAGAAACAAACAGAGAGAGTGAAGTGGCCCTGATCTGCGGC 453
 QY 586 AACAGGCTGAC---GGGACTTTCACCGGAGGAGGAGACCGCGAGATGAGCAGCTG 642
 Db 454 AACAGAGAGAGACAGCGGAGGAGTGTGCGCGAGGTGCGCACCGAGCGCGAGCTGCTG 513
 QY 643 GTGGGAGAGACCGCGAGCGCTGCGCTTACATCTGAGATCTGCGCAAGAAAGAGCAGC 702
 Db 514 GTGTGCGGCGAGC---AGACTGCGGCTTACTTTCGAGGTGTGCGCAAGAAACAGCAAC 570
 QY 703 CTGAGCAAGATGTCGCGCGCGCTCTTGGCATGTGCGCAAGCTTGGCCAGAGAGAGCCCA 762
 Db 571 GTGAGCAAGATGTCGCGCGCGCTCTTTCAGCAATGCGCAAGCTTGGCCAGAGAGAGCCCA 630
 QY 763 GACCTGACCGCAAGATGTCGCGCGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 822
 Db 631 GCGCTGACCGCAAGATGTCGCGCGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 683
 QY 823 AACCAAGAGATGTCGCTT 882
 Db 684 -----CTTTCGATGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTT 720
 QY 883 GGCATGTCGAGCACTTTCGCTT 942
 Db 721 GGCATGTCGTCGCTT 780
 QY 943 GAGAGCGCGAGCTT 993
 Db 781 GCCAAGGTCTTTCGCGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTT 831

RESULT 12

ABL92076 standard; cDNA; 2832 BP.

ID ABL92076;

AC ABL92076;

DE 30-MAY-2002 (first entry)

Human Tumour Endothelial Marker polynucleotide SEQ ID NO 174.

Human; mouse; rat; TEM: tumour endothelial marker; NEM; PEM; cytosolic; normal endothelial marker; pan-endothelial marker; immunostimulant; antiangiogenic; tumour; neovascularisation; vascularised tumour; polycystic kidney disease; diabetes; retinopathy; rheumatoid arthritis; psoriasis; gene; ss.

OS Homo sapiens.

PN WO200210217-A2.

PD 07-FEB-2002.

PF 01-AUG-2001; 2001WO-US24031.

PR 02-AUG-2000; 2000US-222599P.

PR 11-AUG-2000; 2000US-224360P.

PR 11-APR-2001; 2001US-282850P.

PA (UVD) UNIV JOHNS HOPKINS.

PI St Croix B, Kinzler KW, Vogelstein B;

DR WPI: 2002-291856/33.

An isolated molecule comprising an antibody variable region which specifically binds to an extracellular domain of a tumor endothelial marker (TEM) protein, useful for inhibiting tumor growth -

Disclosure: Page 119-120; 331pp: English.

The invention relates to an isolated molecule comprising an antibody variable region which specifically binds to an extracellular domain of a tumor endothelial marker (TEM) protein selected from ABB90732, ABB90740, ABB90749, ABB90750 and ABB90759. The antibodies which bind to TEM proteins have cytostatic, immunostimulant and antiangiogenic activity. They are useful for inhibiting tumor growth, neovascularisation in subjects bearing a vascularised tumour, polycystic kidney disease, diabetic retinopathy, rheumatoid arthritis and psoriasis. Human, mouse and rat TEM genes and the encoded proteins (ABL92075-ABL92141 and ABB90721-ABB90789) are disclosed, as are marker oligonucleotide sequences: tumour endothelial markers (TEM) ABL91996-ABL92041 and ABL92143-ABL92191; normal endothelial markers (NEM) ABL92042-ABL92074; and pan-endothelial markers (PEM) ABL91903-ABL91995.

Sequence 2832 BP; 602 A; 891 C; 727 G; 612 T; 0 other;

Query Match 22 6%; Score 407.4; DB 24; Length 2832;

Best Local Similarity 71.5%; Pred. No. 1.8e-47;

Matches 594; Conservative 0; Mismatches 201; Indels 36; Gaps 3;

166 GCGATGATCAAGAGATGTCGCCGAGGAGCTGAGTATCCCGGCAAGAACTGC 225

34 GCGATGATGAAGACTTGTCCAGGAGGAGTGCACGCTGAGCCGCAAAACTCA 93

226 TATGCAATGTCATCTCGCTGCTCCAGAGTGGCAAGAGGCGCATGTCGCGCTTC 285

94 TACGCGATGTCGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 153

286 CTCACGCGCGGCTTCGAGAGCTTACACGCTTACACGCTTACACGCTTACACGCTT 345

154 CTCATGCGCGCTTTCGAGAGCTTACACGCTTACACGCTTACACGCTTACACGCTT 213

346 TATCCATCCGCGGAGGCTTACACGCTGAGATCTGAGACGTCGCGCAACCCG 405

214 TACAACTCCGCGGAGGCTTACACGCTGAGATCTGAGACGTCGCGCAACCCG 273

406 TTCCCGCGGAGGCTTACACGCTGAGATCTGAGACGTCGCGCAACCCG 465

274 TTCCCGCGGAGGCTTACACGCTGAGATCTGAGACGTCGCGCAACCCG 333

466 CTGAGACGCGGAGGCTTACACGCTGAGATCTGAGACGTCGCGCAACCCG 525

334 CTGAGATACCGGAGGCTTACACGCTGAGATCTGAGACGTCGCGCAACCCG 393

526 AAGCTTCCCTCAAGAACAAACAGAGACGTCGCGCAACCCG 585

394 AAGCTTCCCTCAAGAACAAACAGAGACGTCGCGCAACCCG 453

586 AACAGGCTGAC---CGGCTTACACGCTGAGATCTGAGACGTCGCGCAACCCG 642

454 AACAGGCTGAC---CGGCTTACACGCTGAGATCTGAGACGTCGCGCAACCCG 513

643 GTGGCGGAGGCTTACACGCTGAGATCTGAGACGTCGCGCAACCCG 702

514 GTGGCGGAGGCTTACACGCTGAGATCTGAGACGTCGCGCAACCCG 570

703 CTGAGACGATGTCGCGGCTTACACGCTGAGATCTGAGACGTCGCGCAACCCG 762

571 GTGGACGAGATGTCGCGGCTTACACGCTGAGATCTGAGACGTCGCGCAACCCG 630

763 GACCTGACCGGAGGCTTACACGCTGAGATCTGAGACGTCGCGCAACCCG 822

631 GACCTGACCGGAGGCTTACACGCTGAGATCTGAGACGTCGCGCAACCCG 683

823 AACAGGCTGTCGCGGCTTACACGCTGAGATCTGAGACGTCGCGCAACCCG 882

684 -----CTGTCATGCGCGGCTTACACGCTGAGATCTGAGACGTCGCGCAACCCG 720

883 GGCATGTCGCGGCTTACACGCTGAGATCTGAGACGTCGCGCAACCCG 942

721 GGCATGTCGCGGCTTACACGCTGAGATCTGAGACGTCGCGCAACCCG 780

943 GAGAGGCTGTCGCGGCTTACACGCTGAGATCTGAGACGTCGCGCAACCCG 993

781 GGCATGTCGTCGCGGCTTACACGCTGAGATCTGAGACGTCGCGCAACCCG 831

RESULT 13

ABQ28320/C

ID ABQ28320 standard; DNA; 624 BP.

ABQ28320;

12-JUL-2002 (first entry)

Oligonucleotide for detecting cytosine methylation SEQ ID NO 14911.

Human; cytosine methylation; 5'-CpG-3'; uretic; cytosine; diagnosis;

KW drug; side effect; cancer; central nervous system; cardiovascular;

KW gastrointestinal; respiratory system; single nucleotide polymorphism;

SNP; cell differentiation; ds.

OS Homo sapiens.

PN WO200218632-A2.

07-MAR-2002.

01-SEP-2001; 2001WO-EP10074.

01-SEP-2000; 2000DE-1043826.

05-SEP-2000; 2000DE-1044543.

PA (EPIC-) EPIGENOMICS AG.

XX Olek A, Piepenbrock C, Berlin K, Guetig D;
XX WPI: 2002-371829/40.

XX Determining the degree of cytosine methylation in genomic DNA, useful
XX for diagnosis and prognosis, comprises selective hybridization of
XX amplicons from chemically treated DNA

PS Claim 12: 56pp + Sequence Listing; 56pp; German.

CC This invention describes a novel method for determining the degree of
CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
CC genomic sample of DNA. The sample is treated chemically to convert
CC cytosine (C) but not methylated C, to uracil, then part of the genomic
CC DNA that contains the target C is amplified to form a labeled amplicon.
CC The amplicon is hybridised to two classes, each with at least one
CC member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers
CC and the degree of hybridisation to both classes is determined from the
CC label on the amplicon. From the ratio of labels hybridised to the two
CC classes of oligomers, the degree of methylation is calculated. The method
CC is used: (i) for diagnosis and/or prognosis of side effects of
CC therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders
CC of the central nervous, cardiovascular, gastrointestinal and respiratory
CC systems etc., particularly by detecting mutations or single nucleotide
CC polymorphisms (SNP/s); and (ii) for differentiation of cell or tissue
CC types and for investigating cell differentiation. The method allows the
CC methylation status of many C residues to be determined simultaneously.
CC AB013410-AB054121 represent genomic DNA sequences used to illustrate the
CC method for determining the degree of cytosine methylation described in
CC the disclosure of the invention.

XX Sequence 624 BP: 79 A; 68 C; 212 G; 265 T; 0 other;

XX Query Match 21.7%; Score 391.4; DB 24; Length 624;
XX Best Local Similarity 77.9%; Pred. No. 2.8e-45;
XX Matches 484; Conservative 0; Mismatches 136; Indels 1; Gaps 1;

QY 594 TGACGCGGACTTCTTACCGAGGTGGACGAGATCGAGCAGTGGGGGAGCA 653
DB 620 TAACCGGACTTCTTACCGGAAATTAACCAAGCAATCGAACTTAATTAACGAC 561
QY 654 CCGCCGCGGCGCTTCTTACCGGAGTCTGGCCAAAGAACAGCAGCTGACAGAT 713
DB 560 CCGCCGCGGCGCTTCTTACCGGAGTCTGGCCAAAGAACAGCAGCTGACAGAT 501
QY 714 GTTCCGCGGCGCTTCTTACCGGAGTCTGGCCAAAGAACAGCAGCTGACAGAT 773
DB 500 ATTCCGCGGCGCTTCTTACCGGAGTCTGGCCAAAGAACAGCAGCTGACAGAT 441
QY 774 CAAAGTCTGCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCT 833
DB 440 CAAAGTCTGCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCT 381
QY 834 GCTGGGCGGCGGCGCTTCTTACCGGAGTCTGGCCAAAGAACAGCAGCTGACAGAT 893
DB 380 ACTAGAACCGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 322
QY 894 ACCCTTCCGCGGCGCTTCTTACCGGAGTCTGGCCAAAGAACAGCAGCTGACAGAT 953
DB 321 ACCCTTCCGCGGCGCTTCTTACCGGAGTCTGGCCAAAGAACAGCAGCTGACAGAT 262
QY 954 GCGCGGCGGCGGCGCTTCTTACCGGAGTCTGGCCAAAGAACAGCAGCTGACAGAT 1013
DB 261 GCGCGGCGGCGGCGCTTCTTACCGGAGTCTGGCCAAAGAACAGCAGCTGACAGAT 202
QY 1014 TGGCGACACAACTTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1073
DB 201 TAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 142
QY 1074 CAGGCGGCGGCGGCGCTTCTTACCGGAGTCTGGCCAAAGAACAGCAGCTGACAGAT 1133
DB 141 CAAACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 82

QY 1134 GGAGCGGCGGCGCTTCTTACCGGAGTCTGGCCAAAGAACAGCAGCTGACAGAT 1193
DB 81 AAAACCGGCGGCGCTTCTTACCGGAGTCTGGCCAAAGAACAGCAGCTGACAGAT 22
QY 1194 AAGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1214
DB 21 AAAACTAAACTATTTCCCGAC 1

RESULT 14

ID AB028321 standard; DNA: 624 BP.

AC AB028321;

DT 12-JUL-2002 (first entry)

XX Oligonucleotide for detecting cytosine methylation SEQ ID NO 14912.

KW Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;

KW drug; side effect; cancer; central nervous system; cardiovascular;

KW gastrointestinal; respiratory system; single nucleotide polymorphism;

XX SNP; cell differentiation; ds.

XX Homo sapiens.

XX WO200210632-A2.

XX 01-SEP-2001; 2001WO-EP10074.

XX 01-SEP-2000; 2000DE-1043826.

XX 05-SEP-2000; 2000DE-1044543.

XX (EPIC-) EPIDENOMICS AG.

XX Olek A, Piepenbrock C, Berlin K, Guetig D;

XX WPI: 2002-371829/40.

XX Determining the degree of cytosine methylation in genomic DNA, useful

XX for diagnosis and prognosis, comprises selective hybridization of

XX amplicons from chemically treated DNA

XX Claim 12: 56pp + Sequence Listing; 56pp; German.

XX This invention describes a novel method for determining the degree of

XX methylation of a particular cytosine in a motif 5'-CpG-3', present in a

XX genomic sample of DNA. The sample is treated chemically to convert

XX cytosine (C) but not methylated C, to uracil, then part of the genomic

XX DNA that contains the target C is amplified to form a labeled amplicon.

XX The amplicon is hybridised to two classes, each with at least one

XX member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers

XX and the degree of hybridisation to both classes is determined from the

XX label on the amplicon. From the ratio of labels hybridised to the two

XX classes of oligomers, the degree of methylation is calculated. The method

XX is used: (i) for diagnosis and/or prognosis of side effects of

XX therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders

XX of the central nervous, cardiovascular, gastrointestinal and respiratory

XX systems etc., particularly by detecting mutations or single nucleotide

XX polymorphisms (SNP/s); and (ii) for differentiation of cell or tissue

XX types and for investigating cell differentiation. The method allows the

XX methylation status of many C residues to be determined simultaneously.

XX AB013410-AB054121 represent genomic DNA sequences used to illustrate the

XX method for determining the degree of cytosine methylation described in

XX the disclosure of the invention.

XX Sequence 624 BP: 265 A; 212 C; 68 G; 79 T; 0 other;

XX Query Match 21.7%; Score 391.4; DB 24; Length 624;

XX Best Local Similarity 77.9%; Pred. No. 2.8e-45;

[illegible]

(EPiG-) EPiGENOMICS AG.
Olek A., Piepenbrock C., Berlin K., Guetig D;
WPI: 2002-371829/40.
Determining the degree of cytosine methylation in genomic DNA, useful
for diagnosis and prognosis, comprises selective hybridization of
amplicons from chemically treated DNA -
Claim 12; 56pp + Sequence Listing; 56pp; German.

This invention describes a novel method for determining the degree of methylation of a particular cytosine in a motif 5'-CGG-3', present in a genomic sample of DNA. The sample is treated chemically to convert cytosine (C) but not methylated C, to uracil, then part of the genomic DNA that contains the target C is amplified to form a labeled amplicon. The amplicon is hybridised to two classes, each with at least one member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the degree of hybridisation to both classes is determined from the label on the amplicon. From the ratio of labels hybridised to the two classes of oligomers, the degree of methylation is calculated. The method is used: (i) for diagnosis and/or prognosis of side effects of therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders of the central nervous, cardiovascular, gastrointestinal and respiratory systems etc., particularly by detecting mutations or single nucleotide polymorphisms (SNP's); and (ii) for differentiation of cell or tissue types and for investigating cell differentiation. The method allows the methylation status of many C residues to be determined simultaneously. AB014410-AB054121 represent genomic DNA sequences used to illustrate the method for determining the degree of cytosine methylation described in the disclosure of the invention.

Sequence 624 BP; 127 A; 68 C; 206 G; 223 T; 0 other;

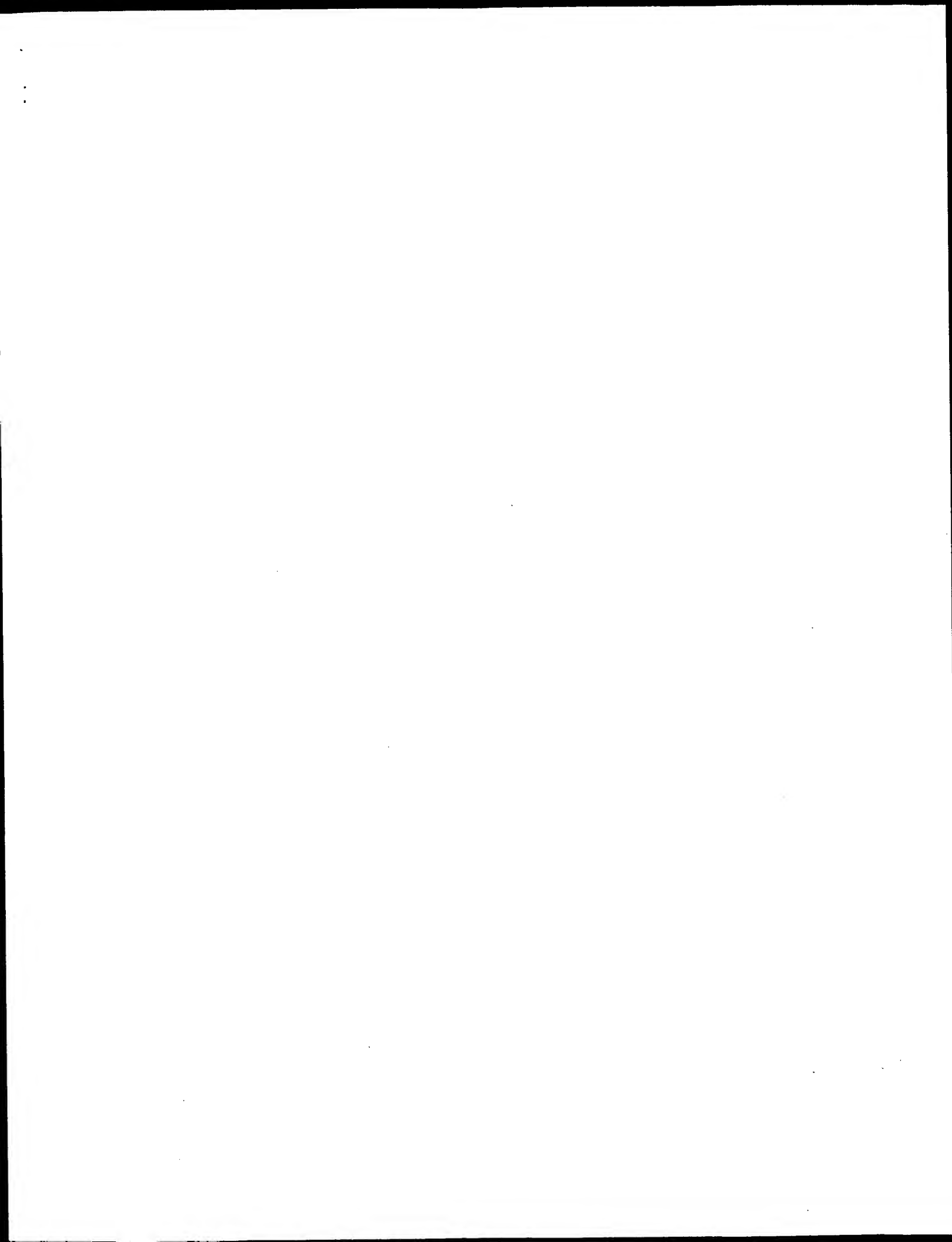
every Match	21.38; Score 383; DB 24; Length 624;
-------------	--------------------------------------

Local Similarity	77.28;	Pred. No. 3.9e-44;
Matches	478;	Conservative
	0;	Mismatches
	140;	Indels
	1;	Gaps
	1	

[illegible]

Db 480 GTTTAGGTCGGAGCGCGCGGATGCGCTTTTTCGGCGATTGTTTTAGTA 539
Qy 1130 CTGGGAGCGCGCCACTGAACGAGAGGAGCGTCATCTCCGAGGAAGAACG 1189
|||
Db 540 TTGGGAGCGCGCTTATTGAATCGAAGGAGCGTTATTGTTGGAAGAAAGAACG 599
|||
Qy 1190 GGCCAAAGACTGGGACTATT 1208
|||
Db 600 GGTTAGATTGGGATTATT 618

Search completed: December 28, 2002, 03:00:10
Job time : 388.134 secs



Query Match	22.8%	Score 409.8	DB 10	Length 3082
Best Local Similarity	66.3%	Pred. No. 6,8e-54		
Matches 654	Conservative 0	Mismatches 297	Indels 36	Gaps 3
QY 10	AGCGGAGCCGGAGACCCCAAGCCCGAGCCGCGCCGACGCCGAGAGAGCCCTCCAGCGGCT	69		
Db	52	AGCGGAGGGGAGAGATCCCGCCAGTGAACCCGGAGAGCCACCAAGACTCTGGGAGGCTCG	111	
QY 70	CACCCGCGCTGCCACCCCAAGCAGCCCTAAGCCGCTCTCTGCGCTTCCTGGCCCGCGCG	129		
Db	112	GCGGCTGGGAGCAGCAGAGCGACGCTCCCGGAGCTCCGGGCTTCCAGGACAGCTCTGAGG	171	
QY 130	CCGCGCTTCGCGGCCCTCTGCCCAATGAAACGGGCGCGGATGATCAGAGATAGTGCCGG	189		
Db	172	CGTGCCAGAGGGCCCGGCCCGCCATTTCCAGGCCCGGAGCGCATGATGAAGACTTTGTCCAGC	231	
QY 190	AGGCACTCGGACTAGATATCCCGGCAAGACTGCTATGCGATGGTCAATCTCGGCGCTGG	249		
Db	232	GGGAGATGCACCTCAAGTGTGCCGCCCAAAATCATACCGCATGGTGGTGGGCTGCC	291	
QY 250	TCCAAAGTGGGCAAGAGGCGCATGCTGTGCGGCTTCTCAGCGCGGCTTTCGAGAGCGCC	309		

```

Db 292 TCTGGGTGGGCAAGCTCCATGTCGTCTGCTTCTCAATGGCCGCTTTGAGACACAG 351
OY 310 TATACGGCTTACCTGAGAGCTTCCACCCGAAATTTCTACTCCATCCGCGAGGTCTTAC 369
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 352 TACACACCCACCTCAGAGAGCTTCCACCCGAAATTTACAAATCTCGCGCGACATGTAC 411
OY 370 CAGCTGACATCTCTGACACGTCCGCGAACACCCGTTCCCGCCATCGGGCGCTCTCC 429
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 412 CAGCTGACATCTCTGACACGTCCGCGAACACCCGTTCCCGCCATCGGGCGCTCTCC 471
OY 430 ATCTCTCAGAGAGAGCTTCTCATCTGATGATGATGATGATGATGATGATGATGAT 489
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 472 ATCTCTCAGAGAGAGCTTCTCATCTGATGATGATGATGATGATGATGATGATGAT 531
OY 490 GAGGTGACAGAGCTTCTCATCTGATGATGATGATGATGATGATGATGATGATGAT 549
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 532 GAGGTGACAGAGCTTCTCATCTGATGATGATGATGATGATGATGATGATGATGAT 591
OY 550 AAGGAGAGAGCTTCTCATCTGATGATGATGATGATGATGATGATGATGATGATG 606
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 592 AAGGAGAGAGCTTCTCATCTGATGATGATGATGATGATGATGATGATGATGATG 651
OY 607 TACGCGAGAGAGCTTCTCATCTGATGATGATGATGATGATGATGATGATGATGATG 666
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 652 TCGCGCAGAGAGCTTCTCATCTGATGATGATGATGATGATGATGATGATGATGATG 708
OY 667 GCTTACTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 726
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 709 GCTTACTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 768
OY 727 TTTGCGATGCGCAAGCTTCTCATCTGATGATGATGATGATGATGATGATGATGATG 786
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 769 TTTGCGATGCGCAAGCTTCTCATCTGATGATGATGATGATGATGATGATGATGATG 828
OY 787 CAGTACTGCGAGCTTCTCATCTGATGATGATGATGATGATGATGATGATGATGATG 846
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 829 CAGTACTGCGAGCTTCTCATCTGATGATGATGATGATGATGATGATGATGATGATG 858
OY 847 AGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 906
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 859 TTTGCGATGCGCAAGCTTCTCATCTGATGATGATGATGATGATGATGATGATGATG 918
OY 907 CGGCCAGAGCTTCTCATCTGATGATGATGATGATGATGATGATGATGATGATGATG 966
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 919 CGGCCAGAGCTTCTCATCTGATGATGATGATGATGATGATGATGATGATGATGATG 978
OY 967 GCCAAGGACAGAGAGCTTCTCATCTGATGATGATGATGATGATGATGATGATGATG 993
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 979 GCCCGTGAAGAGGACAGAGCTTCTCATCTGATGATGATGATGATGATGATGATG 1005
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

RESULT 2
US-09-880-107-2890/c
Sequence 2890, Application US/09880107
Patent No. US20020142981A1

GENERAL INFORMATION:

APPLICANT: Horne, Darci T.
APPLICANT: Vockley, Joseph G.
APPLICANT: Scherf, Uwe

APPLICANT: Gene Logic, Inc.
TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer

FILE REFERENCE: 44921-5028-WO
CURRENT APPLICATION NUMBER: US/09/880,107

CURRENT FILING DATE: 2001-06-14
PRIOR APPLICATION NUMBER: US 60/211,379

PRIOR FILING DATE: 2000-06-14
PRIOR APPLICATION NUMBER: US 60/237,054

PRIOR FILING DATE: 2000-10-02
NUMBER OF SEQ ID NOS: 3950

SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2890

LENGTH: 414
TYPE: DNA

```

; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 R51831
; NAME/KEY: unsure
; LOCATION: (1)..(414)
; OTHER INFORMATION: n = a or c or g or t
US-09-880-107-2890

```

Query Match 19.6%; Score 352.6; DB 10; Length 414;
Best Local Similarity 94.7%; Pred. No. 2,1e-45;
Matches 392; Conservative 0; Mismatches 18; Indels 4; Gaps 3;

```

OY 1284 ATCTTCTCAAGAGCTTGAAGTG-AGCGGGGGTGGGGAGAGATGTG-AACTATCCAG 1341
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 414 AACTTCTCAAGAGCTTGAAGTG-AGCGGGGGTGGGGAGAGATGTG-AACTATCCAG 355
OY 1342 CCTCTGCTAGGCTTGAAGAGACCGTCATGCCC--GCTTGAAGGCTGAGACCGACGGGCA 1399
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 354 CCTCTGCTAGGCTTGAAGAGACCGTCATGCCC--GCTTGAAGGCTGAGACCGACGGGCA 295
OY 1400 TTTATCTTGTCTGTGATTCGCGGTGCTGTGACAGCGCGGTAGAGCCTTGCCTCCGAAA 1459
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 294 TTTATCTTGTCTGTGATTCGCGGTGCTGTGACAGCGCGGTAGAGCCTTGCCTCCGAAA 235
OY 1460 CTAAAGCGGGGGGCGGTGCTGTAATCATAGCCAACTGTTTACATGTGAGTGAAC 1519
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 234 CTAAAGCGGGGGGCGGTGCTGTAATCATAGCCAACTGTTTACATGTGAGTGAAC 175
OY 1520 TGCACAAAGGAACACAAACAAACCTTGCACCTTAAAGCTTAAAGCTTAAAGCTGA 1579
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 174 TGCACAAAGGAACACAAACAAACCTTGCACCTTAAAGCTTAAAGCTTAAAGCTGA 115
OY 1580 CACGAACAAACCTTACCAGGCTTAACTGTTTAACTGTTGTTGAGAGCTTAAAGTTAT 1639
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 114 CACGAACAAACCTTACCAGGCTTAACTGTTTAACTGTTGTTGAGAGCTTAAAGTTAT 55
OY 1640 GCTTATTTGTTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAA 1693
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 54 GCTTATTTGTTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAA 1

```

RESULT 3
US-09-960-352-10273

Sequence 10273, Application US/09960352
Patent No. US20020137139A1

GENERAL INFORMATION:

APPLICANT: Warren, Wesley C.
APPLICANT: Tao, Nengding

APPLICANT: Byatt, John C.
APPLICANT: Mathalagan, Nagappan

TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACRATION AND

FILE REFERENCE: 16511.006/37-21(10298)C
CURRENT APPLICATION NUMBER: US/09/960,352

CURRENT FILING DATE: 2001-09-24
NUMBER OF SEQ ID NOS: 15112

SEQ ID NO 10273
LENGTH: 405

TYPE: DNA
ORGANISM: Bos taurus

FEATURE:
NAME/KEY: unsure

LOCATION: (3)..(31)
OTHER INFORMATION: unsure at all n locations

OTHER INFORMATION: Clone ID: 44-LIB3058-050-Q1-K1-C8
US-09-960-352-10273

Query Match 17.0%; Score 305.8; DB 10; Length 405;
Best Local Similarity 87.3%; Pred. No. 2.5e-38;
Matches 345; Conservative 0; Mismatches 49; Indels 1; Gaps 1;

```

OY 94 CTTACGCGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 152
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

Db 11 CCCAACACACGNCCTGCTGNCCTACAGGCCGCCGCCGCCCTCTGAGCCCTCTGCTC 70
 QY 153 AATGAACATGCGCCGATATCAAGATATGTCGCCGAGCCGATCGAGATGATATCC 212
 Db 71 AATGAACATGTCGCGATATCAAGATATGTCGCCGAGCCGATCGAGATGATATCC 130
 QY 213 GGCAGAACTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 272
 Db 131 GGCAGAACTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 190
 QY 273 CGTGTGCGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 332
 Db 191 CGTGTGCGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 250
 QY 333 CCACCGCAGATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 392
 Db 251 CCACCGCAGATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 310
 QY 393 CGGCAACACCGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 452
 Db 311 CGGCAACACCGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 370
 QY 453 CCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 487
 Db 371 CCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 405

RESULT 4

US-09-778-963A-3
 ; Sequence 3, Application US/09778963A
 ; Patent No. US2002011512A1
 ; GENERAL INFORMATION:
 ; APPLICANT: NEELAM, Beena et al
 ; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,
 ; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE
 ; FILE REFERENCE: CLO01112
 ; CURRENT APPLICATION NUMBER: US/09/778,963A
 ; NUMBER OF SEQ ID NOS: 5
 ; SOFTWARE: FASTSEQ for Windows Version 4.0
 ; SEQ ID NO 3
 ; LENGTH: 11221
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-778-963A-3

Query Match 13.3%; Score 240.2; DB 10; Length 11221;
 Best Local Similarity 68.7%; Pred. No. 3,7e-28;

Matches 392; Conservative 0; Mismatches 143; Indels 36; Gaps 3;

QY 426 CTCCATCTCTACAGAGACGTTTTCATCTGTTGTTTCTGATCTGACACACCGGACCTCTT 485
 Db 7680 CTCTCTCTCTGACAGGATGCTTCTTCTGTTTCTGATCTGACACACCGGACCTCTT 7739
 QY 486 CGAGAGAGTGCAGGCTCTGAGCAGATCTGACACCAAGTCTTCTGCTCAAGAACAA 545
 Db 7740 CGATGAGGTCAAGGCTCTGAGCAGATCTGAGGTTCAAGTCTTCTGCTCAAGAACAA 7799
 QY 546 AACCAAGAGAGTGCAGGCTCTGAGCAGATCTGACACCAAGTCTTCTGCTCAAGAACAA 602
 Db 7800 GACCAAGAGAGTGCAGGCTCTGAGCAGATCTGAGGTTCAAGTCTTCTGCTCAAGAACAA 7859
 QY 603 CTCTTACCGAGAGTGCAGGCTCTGAGCAGATCTGAGGTTTCTGAGGTTTCTGAGGTTT 662
 Db 7860 GCTGTGCGCTTCTGAGGCTCTGAGCAGATCTGAGGTTTCTGAGGTTTCTGAGGTTT 7916
 QY 663 CTGTGCTTCTTCTGAGATCTTCTGAGCAGATCTGAGGTTTCTGAGGTTTCTGAGGTTT 722
 Db 7917 CTGTGCTTCTTCTGAGATCTTCTGAGCAGATCTGAGGTTTCTGAGGTTTCTGAGGTTT 7976
 QY 723 GCTCTTCTGAGCAGATCTTCTGAGCAGATCTGAGGTTTCTGAGGTTTCTGAGGTTT 782

Db 7977 GCTCTTACAGATGCGCCAGCTGACACAGATGAGCCCGCCCTGATCGCAAGATCTC 8036
 QY 783 GCTGAGTACTGAGAGTCTGACACAGAGCGCTGCGGAGCAAGAGTCTGCGGCGC 842
 Db 8037 GCTGAGTACTGAGAGTCTGACACAGAGCGCTGCGGAGCAAGAGTCTGCGGCGC 8069
 QY 843 CGGCAACACCGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 902
 Db 8070 ---CTTCTGATGCGCGCGCTTCAAGAGATGAGCCTTATGAGATGATGATGATGATGAT 8126
 QY 903 GCGCGCGCGCGCGCTTCAAGAGATGAGCCTTATGAGATGAGATGAGATGAGATGAGAT 962
 Db 8127 CGCGCGCGCGCGCTTCAAGAGATGAGCCTTATGAGATGAGATGAGATGAGATGAGAT 8186
 QY 963 CCAGGCGCGCGCGCTTCAAGAGATGAGCCTTATGAGATGAGATGAGATGAGATGAGAT 993
 Db 8187 CCAGGCGCGCGCGCTTCAAGAGATGAGCCTTATGAGATGAGATGAGATGAGATGAGAT 8217

RESULT 5

US-09-864-761-21643
 ; Sequence 21643, Application US/09864761
 ; Patent No. US20020048763A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Penn, Sharon G.
 ; APPLICANT: Rank, David R.
 ; APPLICANT: Hanzel, David K.
 ; APPLICANT: Chen, Wensheng
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 ; FILE REFERENCE: Aecm1ca-X-1
 ; CURRENT APPLICATION NUMBER: US/09/864,761
 ; PRIOR APPLICATION NUMBER: 2001-05-23
 ; PRIOR FILING DATE: 2000-02-04
 ; PRIOR APPLICATION NUMBER: US 60/207,456
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: US 09/632,366
 ; PRIOR FILING DATE: 2000-08-03
 ; PRIOR APPLICATION NUMBER: GB 24263,6
 ; PRIOR FILING DATE: 2000-10-04
 ; PRIOR APPLICATION NUMBER: US 60/236,359
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: PCT/US01/00666
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00667
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00664
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00669
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00665
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00668
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00663
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00662
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00661
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00670
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: US 60/234,687
 ; PRIOR FILING DATE: 2000-09-21
 ; PRIOR APPLICATION NUMBER: US 09/608,408
 ; PRIOR FILING DATE: 2000-06-30
 ; PRIOR APPLICATION NUMBER: US 09/774,203
 ; PRIOR FILING DATE: 2001-01-29
 ; NUMBER OF SEQ ID NOS: 49117
 ; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
 ; SEQ ID NO 21643
 ; LENGTH: 368

```

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; INFORMATION: MAP TO AL022334.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.8
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.2
; OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.9
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.9
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3
; OTHER INFORMATION: NT HIT: g111418034, EVALUATE 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: P13856, EVALUATE 8.00e-12
; OTHER INFORMATION: EST_HUMAN HIT: BE389944.1, EVALUATE 0.00e+00
US-09-864-761-21643

```

```

Query Match          9.6%: Score 173.2; DB 10; Length 368;
Best Local Similarity 77.0%; Pred. No. 2.7e-18;
Matches 211: Conservative 0; Mismatches 63; Indels 0; Gaps 0;

```

```

Db 166 GCGATGATCAAGAGATGTCGCCGAGCGACTCGAGCTAGTATCCCGGCCAAGAACTGC 225
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 95 GCGATGATGAAGACTTGTCTCAGCGGAGACTGCACGCTGTCGCCGCCAATACTCA 154
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 226 TATGCAATGTCATCTCGCTGCTGTCAGAGTGGGCAAGCGCCATCGTGGCGCTTC 285
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 155 TACGCAATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 214
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 286 CTCACCGCGCGCTTCGAGAGCGCTTACACGCTTACGAGACTTCACCGCAAGTTC 345
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 215 CTCATATGCGCGCTTGGAGACAGTATACACACCGACATCGAGAGCTTCACCGTAAGTA 274
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 346 TACTTCATCCCGCGGAGGCTTACACAGCTGACATCTCGACAGTCCGCGCAACACCCG 405
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 275 TACAACATCCCGCGGAGCATGACAGCTGACATCTCGATGATCTCGCAACACACCCG 334
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 406 TTCGCCGATGCGCGCGCTTCATCTCAGCAG 439
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 335 TTCGCCGATGCGCGCGCTTCATCTCAGCAG 368
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

RESULT 6

```

US-09-960-352-5830
; Sequence 5830, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Mathialagan, Nagappan
; APPLICANT: Byatt, John C.
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 5830
; LENGTH: 454
; TYPE: DNA
; ORGANISM: Bos taurus
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (181),(244),(385)
; OTHER INFORMATION: unsure at all n locations
; OTHER INFORMATION: Clone ID: 25-LIB3058-057-Q1-K1-G1
US-09-960-352-5830

```

```

Query Match          5.5%: Score 99; DB 10; Length 454;
Best Local Similarity 55.7%; Pred. No. 4.6e-07;
Matches 186: Conservative 0; Mismatches 148; Indels 0; Gaps 0;

```

```

Qy 229 GCGATGATCAATCTCGGCTCTCCAGAGTGGGCAAGACGCCATGTCGCGCTTCCTC 288
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 71 GCGGATGATGATGCTGCTGCTGCGCGGCTGGCAAGAGCGCTGTGGCAAAAGGTGGGTG 130
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 289 ACCGCGCGCTTCGAGAGACGCTTACAGCGCTTACATCGAGAGCTTCACACCGCAATTTTAC 348
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 131 GCGGCAATCTTCGAGAGCGCTTACAGAGCTTCGAGAGCTTCGAGAGCTTCGAGAGCTTC 190
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 349 TCCATCGCGGAGAGTGTACAGCTGACATCTTCGACAGCTCGGCAACACCGCTTC 408
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 191 GCGTCAAGCACAAGCGGCTGCTGACATCAACCGACACACCGGCGGCGCTAC 250
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 409 CCGCGCATGCGCGCGCTTCATCTCAGAGAGAGCTTTTCACTTCTGCTGCTGCTGCTG 468
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 251 CCGGCGCTGACAGCGCTTCGATTCGAGAGGCTGACGCGCTTCATCTGCTTATTCATC 310
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 469 GACACCGCGACTCTCTGAGAGAGTGCAGGCGCTCAGAGAGCTTCGAGAGCTTCGAGAGCT 528
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 311 ACCAGAGAGCAACCTGAGAGAGCTGAGAGCGCTCTTTGAGAGCTGATCCGTCACACTCAA 370
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 529 TCTGCTTCAGAGCAAAACCAAGAGAGAGCTGG 562
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 371 GGTACAAACCGCAGAAAGTCCCGCTCATCTTG 404
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

RESULT 7

```

US-09-960-352-4253
; Sequence 4253, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Mathialagan, Nagappan
; APPLICANT: Byatt, John C.
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 4253
; LENGTH: 452
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 18-LIB34-086-Q1-E1-E5
US-09-960-352-4253

```

```

Query Match          5.4%: Score 97; DB 10; Length 452;
Best Local Similarity 56.4%; Pred. No. 9.3e-07;
Matches 181: Conservative 0; Mismatches 140; Indels 0; Gaps 0;

```

```

Qy 242 TCGGCTGTCAGAGTGGGCAAGAGCGCCATGCTGCGGCTTCACCGCGCGCTTCG 301
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5 TCGGCTGCGCGCGCTGCGGCAAGAGCGCTTCGTCAGAGTGGGCGCGCAACTTC 64
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 302 AGAGAGCTTCACAGCGCTTACAGAGTTCACAGCAAGTTCATCTCATCTCCGCGCGG 361
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 65 GTGAGGCTGACCTGCGACCATCGAAGATACCTACCGCCAGGCGCTAGGCTGAGCCACA 124
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 362 AGGTCTACAGCTGACATCTCGACAGCTCGGCAACACACCGCTTCGCCGCTGCGG 421
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 125 AGGCGGCTGCGCTGACATCAGACAGACACAGCGGTGGCCGCTACCGGCGCTCAGC 184
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 422 GCTCTCATCTTCACAGAGAGCTTTTCACTCTGCTGCTGCTGCTGCTGCTGCTGCTG 481
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 185 GCTTGGCATTTGCGAGGAGTACCGCTTCATCTGCTTATTCATTCACAGAGCAAA 244
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 482 CCTTCGAGAGAGTGCAGCGCTGAGCAGAGATCTCGACACCAAGTCTTGGCTCAAGA 541
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 245 CCTGAGAGAGCTGAGCGCTCTATAGCTGATCCGTCAACAGTAAAGTAAACACCGCG 304
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 542 ACAAACCAAGAGAGAGCTGG 562
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

Db 305 AAAAGTGGCCGCTATCTGG 325

RESULT 8
US-09-765-298A-25

```

; Sequence 25, Application US/09765298A
; Patent No. US20020137017A1
; GENERAL INFORMATION:
; APPLICANT: ARONHEIM, AMI
; TITLE OF INVENTION: METHOD FOR DETECTION OF PROTEIN-PROTEIN INTERACTIONS AND A KIT THERE
; FILE REFERENCE: 108387.01
; CURRENT APPLICATION NUMBER: US/09/765,298A
; PRIOR FILING DATE: 2001-01-22
; PRIOR FILING DATE: 1998-07-22
; PRIOR APPLICATION NUMBER: IL 128017
; PRIOR FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 25
; LENGTH: 351
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-765-298A-25

```

Query Match 5.1%; Score 92.6; DB 10; Length 551;
Best Local Similarity 57.4%; Pred. No. 4.4e-06;
Matches 167; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

```

OY 226 TATGCGATGGTCACTCGGCTGCTCCAGGTGGGCAAGCGGCATCGTGCCTTC 285
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4 TATAGCTGGTGGTGGGGCGCGGTGGGCAAGGTGGGCTGACCATCCAGCTG 63
OY 286 CTCACCGCGCGCTTCGAGAGCGCTTACACGCTTACATCGAGACTTCCACCGAAGTTC 345
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 64 ATCCAGAACCATTTTGGAGCATAGACCCCTATAGAGATTCTTACCGAAGCAG 123
OY 346 TACGCCGATCCGCGGAGGCTTACACGCTTACATCGAGACTTCCACCGAAGTTC 405
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 124 GTGGTCAATGATGGGAGAGAGCTTGTGGACATCCCTGATACCGCGCCAGAGAGAG 183
OY 406 TTCGCCGATGGGCGGCTTCTCATCTCCAGAGAGAGCTTTCATCCGCTTCACT 465
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 184 TACAGCCCATCGGAGACCATGATCGACGCGGAGGCTTCTGCTGTGTTGCC 243
OY 466 CTGACAAACCGGCGCTTCGAGAGAGTGCAGCGGCTCAGGAGCAGATC 516
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 244 ATCAACAACCAACGATCTTTTGAAGACATCCACAGTACAGGAGGAGCAGATC 294

```

RESULT 9
US-10-104-484-1

```

; Sequence 1, Application US/10104484
; Patent No. US2002015056A1
; GENERAL INFORMATION:
; APPLICANT: Guan, Kun-Lian
; APPLICANT: Stewart, Scott
; TITLE OF INVENTION: METHOD OF INHIBITING CANCEROUS CELL PROLIFERATION USING RAS MUTAN
; FILE REFERENCE: UMICH-0010
; CURRENT APPLICATION NUMBER: US/10/104,484
; CURRENT FILING DATE: 2002-03-25
; PRIOR FILING DATE: 2001-03-23
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 570
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Nucleotide sequence of the GDP-bound Ras mutant, RasN17N69.

```

US-10-104-484-1

Query Match 5.1%; Score 92.6; DB 12; Length 570;
Best Local Similarity 57.4%; Pred. No. 4.5e-06;
Matches 167; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

```

OY 226 TATGCGATGGTCACTCGGCTGCTCCAGGTGGGCAAGCGGCATCGTGCCTTC 285
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 10 TATAGCTGGTGGTGGGGCGCGGTGGGCAAGGTGGGCTGACCATCCAGCTG 69
OY 286 CTCACCGCGCGCTTCGAGAGCGCTTACACGCTTACATCGAGACTTCCACCGAAGTTC 345
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 70 ATCCAGAACCATTTTGGAGCATAGACCCCTATAGAGATTCTTACCGGAAGCAG 129
OY 346 TACTCCATCCGCGGAGGCTTACAGAGCTTACATCTGACATCTGACGCGGCAACCGC 405
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 130 GTGGTCAATGATGGGAGAGAGCTGCTGTTGAGACATCTGATACCGCGGAGAGAG 189
OY 406 TTCGCCGATCCGCGGAGGCTTCTCATCTTCACAGAGAGCTTTCATCTGCTGTTCACT 465
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 190 TACAGCGCATGCGGAAACCATGATGCGACCGGAGGCTTCTGCTGTGTTGCC 249
OY 466 CTGACAAACCGGCGCTTCGAGAGAGTGCAGCGGCTCAGGAGCAGATC 516
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 250 ATCAACAACCAACGATCTTTTGAAGACATCCACAGTACAGGAGCAGATC 300

```

RESULT 10
US-10-104-484-3

```

; Sequence 3, Application US/10104484
; Patent No. US2002015056A1
; GENERAL INFORMATION:
; APPLICANT: Guan, Kun-Lian
; APPLICANT: Stewart, Scott
; TITLE OF INVENTION: METHOD OF INHIBITING CANCEROUS CELL PROLIFERATION USING RAS MU
; FILE REFERENCE: UMICH-0010
; CURRENT APPLICATION NUMBER: US/10/104,484
; CURRENT FILING DATE: 2002-03-25
; PRIOR FILING DATE: 2001-03-23
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 570
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Nucleotide sequence of the wild type Ras mutant, RasN17N69.
US-10-104-484-3

```

Query Match 5.1%; Score 92.6; DB 12; Length 570;
Best Local Similarity 57.4%; Pred. No. 4.5e-06;
Matches 167; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

```

OY 226 TATGCGATGGTCACTCGGCTGCTCCAGGTGGGCAAGCGGCATCGTGCCTTC 285
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 10 TATAGCTGGTGGTGGGGCGCGGTGGGCAAGGTGGGCTGACCATCCAGCTG 69
OY 286 CTCACCGCGCGCTTCGAGAGCGCTTACACGCTTACATCGAGACTTCCACCGAAGTTC 345
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 70 ATCCAGAACCATTTTGGAGCATAGACCCCTATAGAGATTCTTACCGGAAGCAG 129
OY 346 TACTCCATCCGCGGAGGCTTACAGAGCTTACATCTGACATCTGACGCGGCAACCGC 405
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 130 GTGGTCAATGATGGGAGAGAGCTGCTGTTGAGACATCTGATACCGCGGAGAGAG 189
OY 406 TTCGCCGATCCGCGGAGGCTTCTCATCTTCACAGAGAGCTTTCATCTGCTGTTCACT 465
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 190 TACAGCGCATGCGGAAACCATGATGCGACCGGAGGCTTCTGCTGTGTTGCC 249
OY 466 CTGACAAACCGGCGCTTCGAGAGAGTGCAGCGGCTCAGGAGCAGATC 516
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```


ORGANISM: Bos taurus
OTHER INFORMATION: Clone ID: 04-LIB3058-049-01-K1-A11
US-09-960-352-836

Query Match
Best Local Similarity 71.3%; Score 88.8; DB 10; Length 373;
Matches 117; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

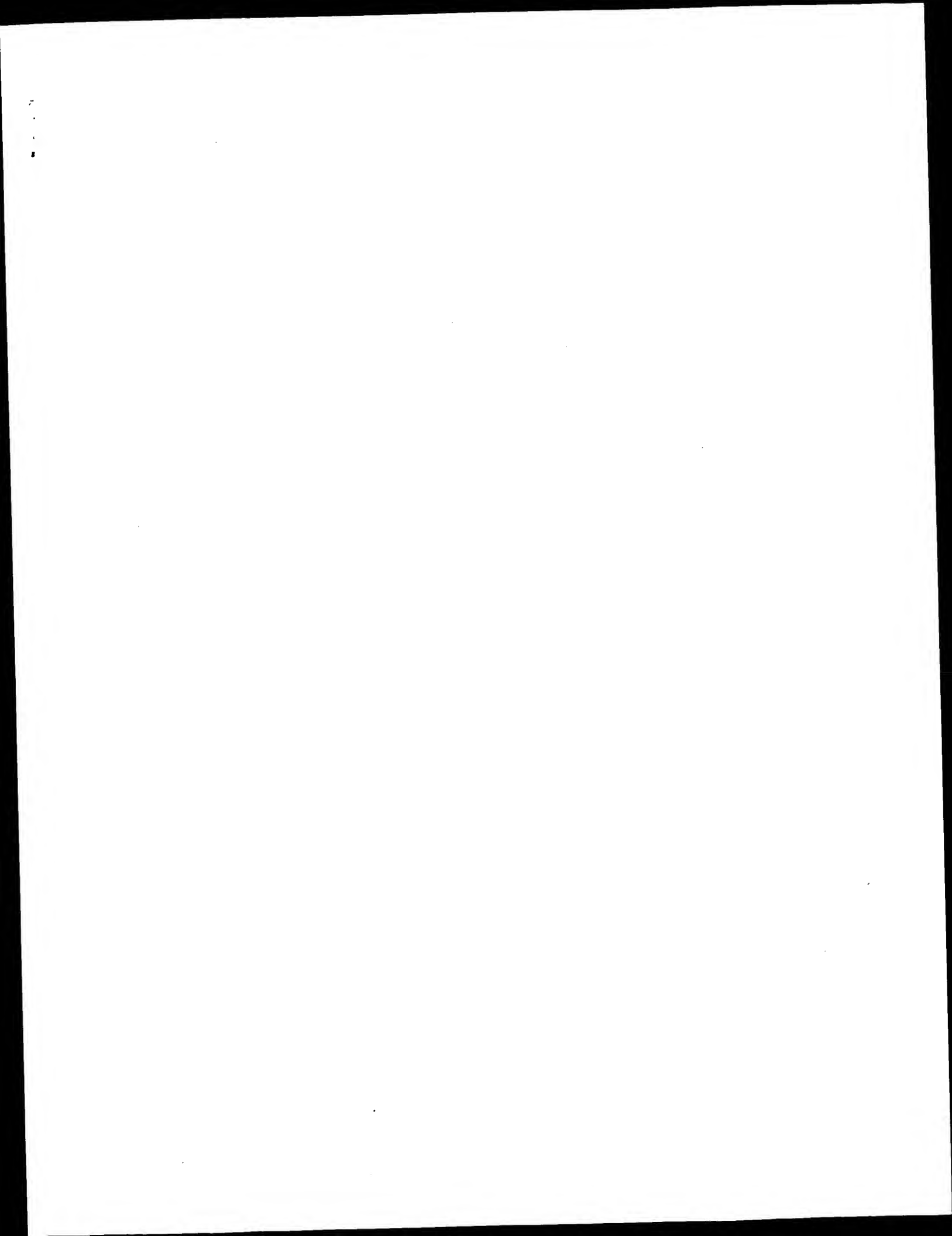
QY 1596 CCCAGCTTTTACTGCTGTGAGCTCTTAAAGTATTCCTTATTTGTTT 1655
DB 173 CCCCGTTTTTTTGTGTTTATTTTATTTTATTTTATTTTGTCCCTTTA 114
QY 1656 TAATATACATTAATTAATTAATGAAAAAATGAAAAAATGAAAAA 1715
DB 113 AAAAAAATGAAAAAATGAAAAAATGAAAAAATGAAAAAATGAAAAA 54
QY 1716 AAAAAAATGAAAAAATGAAAAAATGAAAAAATGAAAAAATGAAAAA 1759
DB 53 AAAAAAATGAAAAAATGAAAAAATGAAAAAATGAAAAAATGAAAAA 10

RESULT 15

US-09-925-301-223
Sequence 223, Application US/09925301
Patent No. US2002005208A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA106
CURRENT APPLICATION NUMBER: US/09/925,301
PRIOR FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05882
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
NUMBER OF SEQ ID NOS: 1694
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 223
LENGTH: 2921
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1609)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (2915)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (2920)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (2921)
OTHER INFORMATION: n equals a,t,g, or c
US-09-925-301-223

Query Match
Best Local Similarity 72.2%; Score 88.8; DB 10; Length 2921;
Matches 114; Conservative 1; Mismatches 43; Indels 0; Gaps 0;

QY 1591 CCTTACCCAGCTTTTACTGCTGTGAGCTCTTAAAGTATTCCTTATTTG 1650
DB 2712 CCTTACCCAGCTTTTACTGCTGTGAGCTCTTAAAGTATTCCTTATTTG 1650
QY 1651 TTTTATATATCAATTAATTAATTAATTAATTAATTAATTAATTAAT 2771
DB 2772 AAGTGTGAAAAAATGAAAAAATGAAAAAATGAAAAAATGAAAAAATG 1710
QY 1711 AAAAAAATGAAAAAATGAAAAAATGAAAAAATGAAAAAATGAAAAA 2831
DB 2832 AAAAAAATGAAAAAATGAAAAAATGAAAAAATGAAAAAATGAAAAA 2869



was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was constructed by Life Technologies. Contact: Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax: (1) 301 610 8371 Email: fliang@lifetech.com URL: http://fulllength.invitrogen.com"

BASE COUNT 168 a 280 c 276 g 200 t 3 others
ORIGIN

Query Match 49.48; Score 889; DB 9; Length 927;
Best Local Similarity 98.88; Pred. No. 5,1e-99;
Matches 914; Conservative 2; Mismatches 7; Indels 2; Gaps 2;

```

OY 725 TCTTGGCCATGCGCAAGCTGCGCCAGCAGATGAGCCAGACCTGACCCGCAAGGTCGCG 784
    |||||||
DB 927 TCTTGGCCATGCGCAAGCTGCGCCAGCAGATGAGCCAGACCTGACCCGCAAGGTCGCG 868
OY 785 TGCAGTACTGCGAGCTGCTGCACAAAGAGCGCTGCCGAAGAAGCTGCTGGCGCGC 844
    |||||||
DB 867 TGCAGTACTGCGAGCTGCTGCACAAAGAGCGCTGCCGAAGAAGCTGCTGGCGCGC 808
OY 845 GCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 904
    |||||||
DB 807 GCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 748
OY 905 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 964
    |||||||
DB 747 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 688
OY 965 AGGCCAAGGACAAGAGGCGCTGCGTACATGAGGAG-CCCGCGCGCGCGCGCGCGCGCA 1023
    |||||||
DB 687 AGGCCAAGGACAAGAGGCGCTGCGTACATGAGGAG-CCCGCGCGCGCGCGCGCGCGCA 628
OY 1024 ACCTAAGGAGGACCTTTTGTAACTCAATCAAGCGCGCGCGCGCGCGCGCGCGCG 1083
    |||||||
DB 627 ACCTAAGGAGGACCTTTTGTAACTCAATCAAGCGCGCGCGCGCGCGCGCGCGCG 568
OY 1084 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1143
    |||||||
DB 567 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 508
OY 1144 CTGAACCGGAGAGGAGCGTCTGCTCCGGAAGGAAGAAAGCGGCGCAAGACTGGGA 1203
    |||||||
DB 507 CTGAACCGGAGAGGAGCGTCTGCTCCGGAAGGAAGAAAGCGGCGCAAGACTGGGA 448
OY 1204 CTATTCCCGACCCCGCGTCCCGCATTTGAGGCCCGCCCATTAATTGAGGAGG 1263
    |||||||
DB 447 CTATTCCCGACCCCGCGTCCCGCATTTGAGGCCCGCCCATTAATTGAGGAGG 388
OY 1264 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1323
    |||||||
DB 387 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 328
OY 1324 GGATGTGAAGTATTCAGAGCTCTGCTAGGCTTCAAGAAACGTCATGCCGCTTGAAGGT 1383
    |||||||
DB 327 GGATGTGAAGTATTCAGAGCTCTGCTAGGCTTCAAGAAACGTCATGCCGCTTGAAGGT 268
OY 1384 CAGGACCCAGGCGGCGTATTTGCTGTGATTCGGGTTGCTGTGACAGCCGGTAGAGC 1443
    |||||||
DB 267 CAGGACCCAGGCGGCGTATTTGCTGTGATTCGGGTTGCTGTGACAGCCGGTAGAGC 208
OY 1444 CTTCTGCGCCCGCGGAACCTAAGGCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1503
    |||||||
DB 207 CTTCTGCGCCCTCCGGAACCTAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 148
OY 1504 TACATGTGAGTGAAGTGCACAAAGGAAACAAACAACTTGAAGCTTAAAGGTAGTT 1563
    |||||||
DB 147 TACATGTGAGTGAAGTGCACAAAGGAAACAAACAACTTGAAGCTTAAAGGTAGTT 88
OY 1564 CCGGTGTCAACATGACACGAAACAAACCTTACCAAGGTGTTTATATCTGTGTGTGTA 1623

```

```

DB 87 CCGGTGTCAACATGACACGAAACAAACCTTACCAAGGTGTTTATATCTGTGTGTA- 29
    |||||||
OY 1624 GGTCTTAAAGTATTCCTTATTT 1648
    |||||||
DB 28 GGTCTTAAAGTATTCCTTATTT 4

```

RESULT 2

BM919341 1035 bp mRNA linear EST 12-MAR-2002
LOCUS
DEFINITION
AGENCOUNT 6715681 NIH_MGC_120 Homo sapiens cDNA clone IMAGE:5748585
5', mRNA sequence.
ACCESSION
BM919341
BM919341.1 GI:19369720
VERSION
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE
1 (bases 1 to 1035)
NIH-MGC http://mgs.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ULNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/ULNL at:
http://image.llnl.gov
Plate: LLM12776 row: j column: 10
High quality sequence stop: 658.

FEATURES

Location/Qualifiers

1..1035
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5748585"
/clone.lib="NIH_MGC_120"
/lab_host="DH10B"
/note="Organ: pooled pancreas and spleen. Vector:
pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of spleen and pancreas from 28 yo
male. Library is oligo-dT primed and directionally cloned
(EcoRV site is destroyed upon cloning). Average insert
size 1.5 kb. Insert size range 1-2.5 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (Invitrogen). Research Genetics
tracking code 025. Note: This is a NIH_MGC Library."

BASE COUNT 209 a 386 c 293 g 146 t 1 others
ORIGIN

Query Match 46.88; Score 842.8; DB 14; Length 1035;
Best Local Similarity 96.98; Pred. No. 1.9e-93;
Matches 882; Conservative 0; Mismatches 22; Indels 6; Gaps 2;

```

OY 7 CCGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 66
    |||||||
DB 40 CCGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 99
OY 67 GCTCACCCTCGGCTCCACCCAGCGAGCTCAGCGCTCTGCGCTTCTCTGCGCCCGG 126
    |||||||
DB 100 GCTCACCCTCGGCTCCACCCAGCGAGCTCAGCGCTCTGCGCTTCTCTGCGCCCGG 159
OY 127 GCGCGCGCTTCGCGCGCGCTTCGCGCAATGAAGTGGCGCGGATGATCAAGAAATGTG 186
    |||||||
DB 160 GCGCGCGCTTCGCGCGCGCTTCGCGCAATGAAGTGGCGCGGATGATCAAGAAATGTG 219
OY 187 CCGAGCGAGCTGAGAGTATCCGCGCAAGAACTGCTATGCTATGCTATGCTGCGG 246
    |||||||
DB 220 CCGAGCGAGCTGAGAGTATCCGCGCAAGAACTGCTATGCTATGCTATGCTGCGG 279

```


Db 737 GCTTACTTGGAGATCTTCGGCCAGAGAACAGACAGCTTGGACAGATGTTCCGCGCTC 796
 QY 727 TTGCGCATGGCCCAAGCTGCGCCAGAGATGAGCCAGACCTGACCGAAGGTCTCGATG 786
 Db 797 TTGCGCATGGCCCAAGCTGCGCCAGAGATGAGCCAGACCTGACCGAAGGTCTCGATG 856
 QY 787 CAGTACTGCGACCTGCTGCAAGAAAGAGCGCTGCG-AGAAACAAGAGCTCTCGCGGCGG 845
 Db 857 CAGTACTGCGACCTGCTGCAAGAAAGAGCGCTGCGGGAACAAGAGCTCTCGCGGCGG 916

RESULT 4
 AL533318 965 bp mRNA linear EST 13-FEB-2001
 LOCUS AL533318 LTI_FL015_Brn1 Homo sapiens cDNA clone CSODN0037019 5

DEFINITION prime, mRNA sequence.
 ACCESSION AL533318
 VERSION AL533318.1 GI:12796811

KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 AUTHORS Li, W. B., Gruber, C., Jesse, J., and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)

COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: segre@genoscope.cns.fr, Web: www.genoscope.cns.fr.
 Location/Qualifiers

FEATURES
 Source
 1..965
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CSODN0037019"
 /clone_1ib="LTI_FL015_Brn1"
 /sex="male"
 /tissue_type="Adult brain"
 /note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was constructed by Life Technologies.
 Contact: Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax: (1) 301 610 8371 Email: filiang@lifestech.com URL: http://fulllength.invitrogen.com"

BASE COUNT 189 a 352 c 273 g 144 t 7 others

Query Match 44.4%; Score 799.2; DB 9; Length 965;
 Best Local Similarity 96.1%; Pred. No. 3.7e-88;
 Matches 855; Conservative 7; Mismatches 20; Indels 8; Gaps 4;

QY 18 CGGACCCCAAGCCGAGCCGCGCCAGAGAGAGCCCTCAGCGCTACCCGCG 77
 Db 55 CGGACCCCAAGCCGAGCCGCGCCAGAGAGAGCCCTCAGCGCTACCCGCG 114
 QY 78 GTGCGACCCCAAGCCGAGCCGCGCTCTGCGCTCTCTGCGCGCGCGCGCTC 137
 Db 115 GTGCGACCCCAAGCCGAGCCGCGCTCTGCGCTCTCTGCGCGCGCGCGCTC 174
 QY 138 GCGGCGCCCTTCCCAATGAACCTGGCGGATGATCAAGAGATGCGCGAGGACTC 197
 Db 175 GCGGCGCCCTTCCCAATGAACCTGGCGGATGATCAAGAGATGCGCGAGGACTC 234
 QY 198 GGAGCTGAGTATCCGGCCCAAGACTGCTATGCTATGCTATGCTGCTGCTCAAGT 257
 Db 235 GGAGCTGAGTATCCGGCCCAAGACTGCTATGCTATGCTATGCTGCTGCTCAAGT 294

QY 258 GGGCAAGACGGCCATCGTGTGCGCTTCTCACCAGCGCGCTTCCAGAGAGCGCTACACGCC 317
 Db 295 GGGCAAGACGGCCATCGTGTGCGCTTCTCACCAGCGCGCTTCCAGAGAGCGCTACACGCC 354
 QY 318 TAACATGAGAGATTTCCACCGCAAGTTCTACTCATCCGCGGCGAGGTCTACACGCTCGA 377
 Db 355 TAACATGAGAGATTTCCACCGCAAGTTCTACTCATCCGCGGCGAGGTCTACACGCTCGA 414
 QY 378 CATCTGCAACAGTTCGGGCAACACCGCTTCCCGGCGATGCGGGGCGCTTCTCATCTCAC 437
 Db 415 CATCTGCAACAGTTCGGGCAACACCGCTTCCCGGCGATGCGGGGCGCTTCTCATCTCAC 473
 QY 438 AGGAGACGTTTTCATCTGTTGTTCACTGTGAGACACCGGAGCTCTTCCAGAGAGTGA 497
 Db 474 AGGAGACGTTTTCATCTGTTGTTCACTGTGAGACACCGGAGCTCTTCCAGAGAGTGA 533
 QY 498 GCGGCTCAGGCGAGATCTTCGACACCAAGCTCTTCCCTCAAGAACCAACCAAGAGAA 557
 Db 534 GCGGCTCAGGCGAGATCTTCGACACCAAGCTCTTCCCTCAAGAACCAACCAAGAGAA 593
 QY 558 CGTGGACGTCCTGCTGATCTGCGGCAACAGAGGTGACCGGACTTTCACCGGAGGT 617
 Db 594 CGTGGACGTCCTGCTGATCTGCGGCAACAGAGGTGACCGGACTTTCACCGGAGGT 653
 QY 618 GGACCAAGCGGAGATCGAGACAGCTGCTGCGGCGAGACACCGGCTGCTGCTTCTGCA 677
 Db 654 GGACCAAGCGGAGATCGAGACAGCTGCTGCGGCGAGACACCGGCTGCTGCTTCTGCA 713
 QY 678 GATCTGCGCAAGAAACAGAGAGCTGAGACAGATGTTCCGCGGCTCTTCCGATGCG 737
 Db 714 GATCTGCGCAAGAAACAGAGAGCTGAGACAGATGTTCCGCGGCTCTTCCGATGCG 773
 QY 738 CAAGCTGCCCAAGAGATGAGAGCCAGACCTGACCGCAAGGCTCTGCTGCTGCTGCA 797
 Db 774 CAAGCTGCCCAAGAGATGAGAGCCAGACCTGACCGCAAGGCTCTGCTGCTGCTGCA 832
 QY 798 CTTGCTGCAAGAAAGAGCGCTTCCGCAAGAAAGAGCTGAGACAGATGTTCCGCGGCT 857
 Db 833 CTTGCTGCAAGAAAGAGCGCTTCCGCAAGAAAGAGCTGAGACAGATGTTCCGCGGCT 886
 QY 858 CGGCGGCGAGCCGCGGAGCGCTTGGCATGCTGAGACCTCTTCCGCGGCG 907
 Db 887 CGGCGGCGAGCCGCGGAGCGCTTGGCATGCTGAGACCTCTTCCGCGGCG 936

RESULT 5
 BM920514 1103 bp mRNA linear EST 12-MAR-2002
 LOCUS BM920514 AGENCOURT 6709473 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5750507
 DEFINITION 5' mRNA sequence.

ACCESSION BM920514
 VERSION BM920514.1 GI:19370893

KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 AUTHORS NIH-MGC http://mgi.ncl.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-r@mail.nih.gov

Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov

Plate: L14M12761 row: j column: 12
 High quality sequence stop: 676.
 Location/Qualifiers

FEATURES

QY	67	GCACCCCGGTGCACACCCAGAGACCCCTCAGCCGCTCTCTGCTCCCTTCTCTGCGGCCG	126
Db	113	GCTACCCCCCGGTGCACACCCAGAGACCCCTCAGCCGCTCTCTGCTCCCTTCTCTGCGGCCG	172
QY	127	CGCCCGCCTTCGCGCGCCCTCTGTGCCAATGAACCTGGCCCGCATGATCAAGAAGATGTGC	186
Db	173	CGCCCGCCTTCGCGCGCCCTCTGTGCCAATGAACCTGGCCCGCATGATCAAGAAGATGTGC	232
QY	187	CCGAGCGACATCGGAGCTGAGATTCGCCGGCCAAAGACGTGATCGCATGAGTCAATCCTCGGC	246
Db	233	CCGAGCGACATCGGAGCTGAGATTCGCCGGCCAAAGACGTGATCGCATGAGTCAATCCTCGGC	292
QY	247	TCGTCCAAAGGTGGGCAAGACGGCCATGTGTGCGCTTCCCTCAACGGGCGCTTTCGAGGAC	306
Db	293	TCGTCCAAAGGTGGGCAAGACGGCCATGTGTGCGGCTTCCCTCAACGGGCGCTTTCGAGGAC	352
QY	307	GCTTACACGCGCTTACCATTCAGAGGACTTCACCCGCNAAGTTTACTTCATCCGCGGCGAGGTC	366
Db	353	GCTTACACGCGCTTACCATTCAGAGGACTTCACCCGCNAAGTTTACTTCATCCGCGGCGAGGTC	412
QY	367	TACACGCTCCGACATCTCTCGACAGTCCGGGACCAACCCGTTTCCCGCGCATGCGGGCCCTTC	426
Db	413	TACACGCTCCGACATCTCTCGACAGTCCGGGACCAACCCGTTTCCCGCGCATGCGGGCCCTTC	472
QY	427	TCCATCCTTCACAGGAGACGTTTTCATCCTGGTGTTCAGTGTGGACAAACGCGACTCTTC	486
Db	473	TCCATCCTTCACAGGAGACGTTTTCATCCTGGTGTTCAGTGTGGACAAACGCGACTCTTC	532
QY	487	GAGGAGGTGCAGCGGCTTCAGGCGACAGATCCTGCACACCAAGTCTTGCTCAAGAAACAA	546
Db	533	GAGGAGGTGCAGCGGCTTCAGGCGACAGATCCTGCACACCAAGTCTTGCTCAAGAAACAA	592
QY	547	ACCAAGGAGAACGTCGAGACGTGCCCTGTCTATCTGCGGGCAACAAGGTCAGCGCGACTTC	606
Db	593	ACCAAGGAGAACGTCGAGACGTGCCCTGTCTATCTGCGGGCAACAAGGTCAGCGCGACTTC	652
QY	607	TACCGCGAGGTGCACAGCGCGAGATGAGACAGCTGGTGGGCGACGACCCCAAGCGCTGC	666
Db	653	TACCGCGAGGTGCACAGCGCGAGATGAGACAGCTGGTGGGCGACGACCCCAAGCGCTGC	712
QY	667	GCTTACTTTCAGATCTCTGGGCGCAGAAACAAACAGCAG--CTGGACCAAGATGTTCCGGCGCT	725
Db	713	GCTTACTTTCAGATCTCTGGGCGCAGAAACAAACAGCAGCTTGGACAGATGTTCCGGCGCT	772
QY	726	CTTGCGCATGAGCAAGCTGCGCCAGACGAGATGAGGCCAGACTGCACCGCAAG--TCTC	782
Db	773	CTTGCGCATGAGCAAGCTGCGCCAGACGAGATGAGGCCAGACTGCACCGCAAGGTTCTCG	832
QY	783	GGTGCAAGTACTGCGAGGTGCTGCACAAGAGAGCGCTGCGAACAAGAACTGCTGCGGGC	842
Db	833	GGTGCAAGTACTGCGAGGTGCTGCACAAGAGAGCGCTGCGAACAAGAAACTGCTGCGG	892
QY	843	CGGCGAGGCGGCGGGGCGGGCGACCCGGGCGGAC	876
Db	893	GGGCGGGAACCGGTGAGCGGGCGGGCGGCGGGAAGAC	926

RESULT	7
BMS543630	
LOCUS	1023 bp mRNA linear EST 20-FEB-2008
DEFINITION	AACNCOURT_6492527 NIH_MGC_124 Homo sapiens cDNA IMAGE:572626
ACCSSION	BMS543630
VERSION	BMS543630.1 GI:18774186
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. 1 (bases 1 to 1023)
TITLE	NIH-MGC http://mgc.nci.nih.gov/ , National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL
COMMENT

JOURNAL
COMMENT
Unpublished (1999)
Contact: Robert Strausberg, Ph. D.
Email: c9apbs-remail.nih.gov
Tissue Procurement: Invitrogen
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/ILNL at:
<http://image.llnl.gov>
Plate: ILAM12719 row: m column: 05
High quality sequence stop: 637.

FEATURES
SOU

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="5726764"
/clone_lib="NH_MGC_124"
/lssue_type="hippocampus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMV-Sport6; Site_1: EcoRV
(destroyed); Site_2: NotI; RNA source male hippocampus,
age 27. Library is oligo-dT primed and directionally
cloned (EcoRV site is destroyed upon cloning). Average
insert size 1.4 kb, insert size range 0.9-4 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (Invitrogen). Research Genetics
tracking code 012."

```

Query Match	40.7%;	Score 733;	DB 13;	Length 1023;
Best Local Similarity	98.3%;	Pred. No. 3.7e-80;		
Matches 762;	Conservative	0;	Mismatches 10;	Indels 3;
				Gaps 2;

QY	7	CCGAGCGAGGCGCGAGGCCCAAGGCCGAGCGCGCGCCCGAGCCGAGCGAGACCCCTCCAGCC	66
Db	2	CCGAGCGGAGCGCGAGGCCCAAGGCCGAGCGCGCGCCCGAGCCCGAGCGAGACCCCTCCAGCC	61
QY	67	GCTACACCCCGGCGTGCACACCCAGGAGACCCCTACGCGCGCTCTCTGCGCCCTCTCTCTGGGCCCC	126
Db	62	GCTACACCCCGGCGTGCACACCCAGGAGACCCCTACGCGCGCTCTCTGCGCCCTCTCTGGGCCCC	121
QY	127	CGCCCGCCCTCGCGCGCCCTCTGCGCCAAATACTGGCGCGCGATGATCAAGAATGTGC	186
Db	122	CGCCCGCCCTCGCGCGCCCTCTGCGCCAAATACTGGCGCGCGATGATCAAGAATGTGC	181
QY	187	CCGAGCGACTCGGAGCTGAGTATCCGGGCCAAGAACTGTAATGCGATGTATCATCTCGGC	246
Db	182	CCGAGCGACTCGGAGCTGAGTATCCCGGCCAAGAACTGTAATGCGATGTATCATCTCGGC	241
QY	247	TGCGTCAAGGTGGGCAAGAGCGCCCATGTGTGCGCGCTTCTCTCAACCGGCGCGCTTCGAGGAC	306
Db	242	TGCGTCAAGGTGGGCAAGAGCGCCCATGTGTGCGCGCTTCTCTCAACCGGCGCGCTTCGAGGAC	301
QY	307	GCTACACGCGCTTACCATTCGAGGACTTCACCGCAAGTTTACTCCATCCCGCGCGAGGTC	366
Db	302	GCTACACGCGCTTACCATTCGAGGACTTCACCGCAAGTTTACTCCATCCCGCGCGAGGTC	361
QY	367	TACCAAGCTCGACATCTTGACACGTCCGGCAACCAACCCGTTCCCGCGCATGCGCGGCTC	426
Db	362	TACCAAGCTCGACATCTTGACACGTCCGGCAACCAACCCGTTCCCGCGCATGCGCGGCTC	421
QY	427	TTCACATCTCAACAGGAGAGGTTTATCTGTGTTCAGTGTGGCAACCGCGACTCTTC	486
Db	422	TTCACATCTCAACAGGAGAGGTTTATCTGTGTTCAGTGTGGCAACCGCGACTCTTC	481
QY	487	GAGGAGGTGCAGCGGCTCAGGCAAGCATCTCTGCACACCAAGTTTGCTCTCAAGAACAA	546
Db	482	GAGGAGGTGCAGCGGCTCAGGCAAGCATCTCTGCACACCAAGTTTGCTCTCAAGAACAA	541
QY	547	ACCAAGGAGAACGTGAGCGTGCCTCTGTGTCATCTGCGGCAACAGGAGTACCGCGACTTC	606
Db	542	ACCAAGGAGAACGTGAGCGTGCCTCTGTGTCATCTGCGGCAACAGGAGTACCGCGACTTC	601

Db 61 GATGTGCCGAGAGCTCGAGAGCTAGATATCCGGCCAAAGACTGATATGATGAT 120
 QY 240 CTTGGCTGTCTCAAGGTGGCAAGAGCGGATCTGTGTCGGCTTCTCAACGGCGCTT 299
 Db 121 CTTGGCTGTCTCAAGGTGGCAAGAGCGGATCTGTGTCGGCTTCTCAACGGCGCTT 180
 QY 300 CGAGGAGCGCTTCAAGGTGGCAAGAGCGGATCTGTGTCGGCTTCTCAACGGCGCTT 180
 Db 181 CGAGGAGCGCTTCAAGGTGGCAAGAGCGGATCTGTGTCGGCTTCTCAACGGCGCTT 240
 QY 360 CGAGGAGCGCTTCAAGGTGGCAAGAGCGGATCTGTGTCGGCTTCTCAACGGCGCTT 419
 Db 241 CGAGGAGCGCTTCAAGGTGGCAAGAGCGGATCTGTGTCGGCTTCTCAACGGCGCTT 300
 QY 420 GGGCCCTTCAAGGTGGCAAGAGCGGATCTGTGTCGGCTTCTCAACGGCGCTT 479
 Db 301 GGGCCCTTCAAGGTGGCAAGAGCGGATCTGTGTCGGCTTCTCAACGGCGCTT 360
 QY 480 CTCCTTCAAGGTGGCAAGAGCGGATCTGTGTCGGCTTCTCAACGGCGCTT 539
 Db 361 CTCCTTCAAGGTGGCAAGAGCGGATCTGTGTCGGCTTCTCAACGGCGCTT 420
 QY 540 GAAACAAACCAAGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 599
 Db 421 GAAACAAACCAAGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 480
 QY 600 GAGCTTCAAGGTGGCAAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 659
 Db 481 GAGCTTCAAGGTGGCAAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 540
 QY 660 GCGCTGGCGCTTCAAGGTGGCAAGAGCTGAGAGCTGAGAGCTGAGAGCT 719
 Db 541 GCGCTGGCGCTTCAAGGTGGCAAGAGCTGAGAGCTGAGAGCTGAGAGCT 600
 QY 720 GCGCTGGCGCTTCAAGGTGGCAAGAGCTGAGAGCTGAGAGCTGAGAGCT 779
 Db 601 GCGCTGGCGCTTCAAGGTGGCAAGAGCTGAGAGCTGAGAGCTGAGAGCT 660
 QY 780 CTCGGTGGCAAGGTGGCAAGAGCTGAGAGCTGAGAGCTGAGAGCT 839
 Db 661 CTCGGTGGCAAGGTGGCAAGAGCTGAGAGCTGAGAGCTGAGAGCT 720
 QY 840 GGGCG 844
 Db 721 GTCGG 725

RESULT 11
 B1596688 814 bp mRNA linear EST 07-SEP-2001
 LOCUS 603243262P1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5285894 5',
 DEFINITION mRNA sequence.
 ACCESSION B1596688
 VERSION B1596688.1 GI:15489627
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 814)
 AUTHORS NIH-MGC <http://mgc.nhl.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-r@mail.nih.gov
 Tissue Procurement: Miklos Palokovits, M.D., Ph.D.
 CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiroki
 Toshiyuki and Piero Carninci (RIKEN)
 DNA Sequencing Arrayed by: The I.M.A.G.E. Consortium (LNLW)
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNLW at:
<http://image.llnl.gov>

Plate: L1M11722 row: k column: 15
 High quality sequence stop: 803.
 Location/Qualifiers
 1. 814
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5285894"
 /clone_lib="NIH_MGC_96"
 /tissue_type="hypothalamus"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: Bluescript (modified
 pluescript KS+); Site.1: BamHI; Site.2: SalI; xhoI (gtcag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTN-3',
 size-selected for average insert size 2.3 kb and
 normalized to R0F 5. This is a primary library enriched
 for full-length clones and constructed using the
 cap-trapper method (Carninci, in preparation). Library
 constructed by M. Brownstein (NHGRI/NHGR). Library
 Institutes of Health). Note: this is a NIH_MGC Library."
 ORIGIN
 BASF COUNT 160 a 310 c 225 g 119 t
 Query Match 37.9%; Score 683.2; DB 13; Length 814;
 Best Local Similarity 98.4%; Pred. No. 4.3e-74;
 Matches 732; Conservative 0; Mismatches 8; Indels 4; Gaps 4;
 QY 7 CCGAGGCGAGCGCGAGAGCGCGAGAGCGCGAGAGCGCGAGAGCGCGAGCGCG 66
 Db 73 CCGAGGCGAGCGCGAGAGCGCGAGAGCGCGAGAGCGCGAGAGCGCGAGCGCG 132
 QY 67 GCTACACCGCGGTGCGACCGAGAGCGCGAGAGCGCGAGAGCGCGAGAGCGCG 126
 Db 133 GCTACACCGCGGTGCGACCGAGAGCGCGAGAGCGCGAGAGCGCGAGAGCGCG 192
 QY 127 GCG 186
 Db 193 GCG 252
 QY 187 CCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCG 246
 Db 253 CCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCG 312
 QY 247 TCGTCAAGGTGGCAAGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAG 306
 Db 313 TCGTCAAGGTGGCAAGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAG 372
 QY 307 GCGTACAGCGCGCTTCAAGGTGGCAAGAGCGAGCGAGCGAGCGAGCGAGCG 366
 Db 373 GCGTACAGCGCGCTTCAAGGTGGCAAGAGCGAGCGAGCGAGCGAGCGAGCG 432
 QY 367 TACCAAGCTGACATCTCGACAGAGCGAGCGAGCGAGCGAGCGAGCGAGCG 426
 Db 433 TACCAAGCTGACATCTCGACAGAGCGAGCGAGCGAGCGAGCGAGCGAGCG 492
 QY 427 TCCATCTTCAAGAGAGCGTTTCAATCTGCTTCAATCTGCTTCAATCTGCTT 486
 Db 493 TCCATCTTCAAGAGAGCGTTTCAATCTGCTTCAATCTGCTTCAATCTGCTT 552
 QY 487 GAGGAGGTGACAGCGGTGACAGCGAGCGAGCGAGCGAGCGAGCGAGCGAG 546
 Db 553 GAGGAGGTGACAGCGGTGACAGCGAGCGAGCGAGCGAGCGAGCGAGCGAG 612
 QY 547 ACCAAGAGAGAGGTGAGAGCGGTGATCTGCGGCAAGAGGTGAGAGCGAGCG 606
 Db 613 ACCAAGAGAGAGGTGAGAGCGGTGATCTGCGGCAAGAGGTGAGAGCGAGCG 672
 QY 607 TACCGGAGGTGAGAGCGAGAGCGAGAGCGAGAGCGAGAGCGAGAGCGAG 665
 Db 673 TACCGGAGGTGAGAGCGAGAGCGAGAGCGAGAGCGAGAGCGAGAGCGAG 732
 QY 666 GCGCTACTTTCGAG-ATCTCGCGCGCAAGAGAGCGAGCGAGCGAGCGAGCG 724
 Db 733 GCGCTACTTTCGAGATCTCGCGCGCAAGAGAGCGAGCGAGCGAGCGAGCG 790

100

Db 491 CTCATCTCTACAGAGAGCGTTTTCATCTGCTGTTCAAGTGTGACACCGGACTCTT 550
 QY 486 CGAGAGGTGACAGCGGCTCAGGACAGACATCTCTGACACCAAGTCTTGGCTCAGAACA 545
 Db 551 CGAGAGGTGACAGCGGCTCAGGACAGACATCTCTGACACCAAGTCTTGGCTCAGAACA 610
 QY 546 AACCAAGAGAGAGCGTGGGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGG 605
 Db 611 AACCAAGAGAGAGCGTGGGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGG 670
 QY 606 CTACCGGCA-GGTGACACAGCGCGA--GATCAGACAGCTGTGGGCGACACCGCGCGG 662
 Db 671 CTACCGGCAAGGTGGACGCGCGAAGATCAGACAGCTGTGGGCGAGCA--CCCGAGCG 729
 QY 663 CTGCGCTACTTCTGAGATCTGGGCGAAGAGACAGAGCGTGGACAGATGTTCCGCGC 722
 Db 730 CTGCGCTACTTCTGAGATCTGGGCGAAGAGACAGAGCGTGGACAGATGTTCCGCGC 788
 QY 723 GCTCTTCCGCAATGCGCAAGCTGCGGCGAGAGATGAGCGCGAGAC 765
 Db 789 GCTCTTCCGCAAGCGCAAGCTG--CCAGCGAGATGAGCGCGAGAC 830

RESULT 15
 B0719566 958 bp mRNA linear EST 16-JUL-2002
 LOCUS B0719566
 DEFINITION IMAGE:6193170 5', mRNA sequence.
 ACCESSION B0719566
 VERSION B0719566.1 GI:21858463
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 1 (bases 1 to 958)
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: sgabds-remail.nih.gov
 Tissue Procurement: Dr. James R. Lupski
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LIML at:
 http://image.liml.gov
 Plate: LLM13596 row: f column: 19
 High quality sequence stop: 514.

FEATURES
 Source
 1..958
 location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:6193170"
 /clone_lib="Lupski_sympathetic_trunk"
 /sex="male"
 /tissue_type="sympathetic trunk"
 /dev_stage="adult, 16 yr"
 /lab_host="DH10B"
 /note="Vector: PCMV-SPORT6 (Life Technologies); site_1:
 Motic; site_2: SalI; cDNA made by oligo-dT priming.
 Directionally cloned using the following adaptors:
 5'-TCGACCCACGCGCG-3' and
 5'-GACTACTTCTAGATCGCGAGCGCGCGCTT(15)-3'. Size selected >
 1 kb for average insert length 1.9 kb. This is a primary
 library, non-amplified. Library constructed by Life
 Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
 College of Medicine); available through Life
 Technologies." 185 a 339 c 275 g 159 t

BASE COUNT
 ORIGIN

Query Match 36.1%; Score 650.8; DB 14; Length 958;
 Best Local Similarity 85.1%; Pred. No. 3,4e-70;
 Matches 781; Conservative 0; Mismatches 117; Indels 20; Gaps 4;

QY 196 TCGGAGCTGAGTATCCCGGCAAGAACTGCTATGCAATGCTATCTGCTGCTCAAG 255
 Db 1 TCGGAGCTGAGTATCCCGGCAAGAACTGCTATGCAATGCTATCTGCTGCTCAAG 60
 QY 256 GTGGGCAAGAGCGGCTATCTGTGGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGG 315
 Db 61 GTGGGCAAGAGCGGCTATCTGTGGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGG 120
 QY 316 CTTACCATGAGAGACTTCCACCGCAAGTTTCTATCTATCCGCGGAGGTCTACAGCTC 375
 Db 121 CTTACCATGAGAGACTTCCACCGCAAGTTTCTATCTATCCGCGGAGGTCTACAGCTC 180
 QY 376 GACATCTCTGACACGCTGCGGCAACACCGCTTCCCGGATGCGGCGCTCTCATCTC 435
 Db 181 GACATCTCTGACACGCTGCGGCAACACCGCTTCCCGGATGCGGCGCTCTCATCTC 240
 QY 436 ACAGAGAGAGTCTTCACTCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 495
 Db 241 ACAGAGAGAGTCTTCACTCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
 QY 496 CAGCGGCTCAGGACAGAGATCTCTGACACCAAGTCTTGGCTCAAGAAACCAAGAGAG 555
 Db 301 CAGCGGCTCAGGACAGAGATCTCTGACACCAAGTCTTGGCTCAAGAAACCAAGAGAG 360
 QY 556 AACGTGAGAGTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 615
 Db 361 AACGTGAGAGTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
 QY 616 GTGGACCAAGCGGAGATGAGAGAGTGTGGGCGACAGACCGGCGGCGGCTGCTGCTGCTGCT 675
 Db 421 GTGGACCAAGCGGAGATGAGAGAGTGTGGGCGACAGACCGGCGGCGGCTGCTGCTGCTGCTG 480
 QY 676 GAGATCTGCGGCAAGAAAGAGAGAGCGCTGAGCAAGTGTGGCGGCGGCTGCTGCTGCTGCTG 735
 Db 481 GAGATCTGCGGCAAGAAAGAGAGAGCGCTGAGCAAGTGTGGCGGCGGCTGCTGCTGCTGCTG 540
 QY 736 GCCAAGCTGCCAGGAGATGAGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 795
 Db 541 GCCAAGCTGCCAGGAGATGAGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 600
 QY 796 GAGCTCTGACCAAGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 850
 Db 601 GAGCTCTGACCAAGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 660
 QY 851 GGCG 907
 Db 661 GGCG 720
 QY 908 GGCG 965
 Db 721 GGCG 780
 QY 966 GGCCAAGAGCAAGAGAGCGGCTGCG-----TCATCAGTGAAGAGCGGCGGCGGCGGCTG 1015
 Db 781 ACCTTGAAGGCGCAAGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 840
 QY 1016 GCGACCAACACTAAGAGAGAGCGGCTTGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1075
 Db 841 TGGGAGCAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 900
 QY 1076 GCGGCGGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 918
 Db 901 TGGGCGGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 918

Search completed: December 28, 2002, 05:50:47
 Job time: 2391.42 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: December 30, 2002, 17:30:01 ; Search time 3130 Seconds
(without alignments)

3086.943 Million cell updates/sec

Title: US-09-709-103-3F1
Sequence: 1 GISSGAGAPSPSAQPEQSP.....IREKASAGSQAKKRCVCIS 332

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 segs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xlp
-Q/cg2.1/USP10.spool/STULLIVANP1/runat.30122002.144154.22137/app.query.fasta_1.519
-DB=GenEmbl -QFWT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=Dlosum62 -TRANS=human40.cdi -LIST=45
-LOCALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-USER=STULLIVANP1 @CG2.1.1 3637 @runat.30122002.144154.22137 -NCPU=6 -ICPU=3
-WAP.TIMEOUT=30 -THREADD=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELDP=6 -DELEXT=7

Database : GenEmbl.*
1: gb_da.*
2: gb_hg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sci.*
12: gb_sy.*
13: gb_un.*
14: gb_vt.*
15: em_da.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sci.*
28: em_un.*

Fig. 3 frame 1

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB ID	Description
1	1721	99.4	1758	9 BC018041	BC018041 Homo sapi
2	1714	99.0	1740	9 AF069506	AF069506 Homo sapi
3	1708	98.7	1746	9 AF177335	AF177335 Homo sapi
4	1659	95.8	4990	9 AF222979	AF222979 Homo sapi
5	1659	95.8	18334	9 AC020558	AC020558 Homo sapi
6	1656	95.7	5141	9 AF262018	AF262018 Homo sapi
7	1627	94.0	58882	2 AC073621	AC073621 Homo sapi
8	1615	93.3	183598	2 AC090608	AC090608 Homo sapi
9	1537	88.8	979	9 AF172846	AF172846 Homo sapi
10	1467	84.7	1623	10 AF009246	AF009246 Mus muscu
11	1453.5	84.0	1612	10 BC034166	BC034166 Mus muscu
12	1449	83.7	1187	9 AF498923	AF498923 Homo sapi
13	1449	83.7	1187	9 AF153192	AF153192 Homo sapi
14	1435.5	82.9	1616	10 AF239157	AF239157 Rattus no
15	1434	82.8	179124	10 AL603710	AL603710 Rattus no
16	1434	82.8	247899	2 AC122995	AC122995 Mus muscu
17	985	56.9	162504	2 AC122995	AC122995 Mus muscu
18	929.5	53.7	2973	6 BC013419	BC013419 Rattus no
19	929.5	53.7	3058	9 BC013419	BC013419 Rattus no
20	898.5	51.9	3469	10 AF134409	AF134409 Homo sapi
21	898	51.9	2832	6 AX393244	AX393244 Sequence
22	867	50.1	2832	9 AF279143	AF279143 Homo sapi
23	867	50.1	3020	6 AX393362	AX393362 Sequence
24	790	45.6	2699	9 HSM809172	HSM809172 Homo sapi
25	526	30.4	114771	9 AC076974	AC076974 Homo sapi
26	519.5	30.0	211071	10 AC076974	AC076974 Homo sapi
27	364.5	21.1	4412	3 AB062937	AB062937 Macaca fa
28	359.5	20.8	181528	3 AC010562	AC010562 Drosophi
29	359.5	20.8	207684	2 AC010562	AC010562 Drosophi
30	359.5	20.8	285495	3 AE003560	AE003560 Drosophi
31	359	20.7	22394	10 BC026377	BC026377 Mus muscu
32	358	20.7	66031	2 AC118046	AC118046 Mus muscu
33	354.5	20.5	4167	6 AX477382	AX477382 Sequence
34	347.5	20.1	177540	9 AC006538	AC006538 Homo sapi
35	340	19.6	2827	9 AK096600	AK096600 Homo sapi
36	340	19.6	3391	9 BC013060	BC013060 Homo sapi
37	336.5	19.4	1410	9 BC013106	BC013106 Homo sapi
38	336	19.4	1774	9 AB076888	AB076888 Homo sapi
39	334.5	19.3	1875	9 BC008065	BC008065 Homo sapi
40	334	19.3	2505	9 AB076889	AB076889 Homo sapi
41	334	19.3	143299	9 AL353619	AL353619 Human DNA
42	334	19.3	143299	9 AC021583	AC021583 Homo sapi
43	334	19.3	196612	9 AC021582	AC021582 Homo sapi
44	331	19.1	2251	10 BC003871	BC003871 Mus muscu
45	329.5	19.0	3061	6 AX430418	AX430418 Sequence

RESULT 1

ALIGNMENTS

BC018041 1758 bp mRNA linear PRI 06-DEC-2001
Homo sapiens, RAS, dexamethasone-induced l, clone MGC:26290
IMAGE:4797482, mRNA, complete cds.
BC018041 GI:17390075
MGC:
Homo sapiens.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1758)
Strausberg, R.
Direct Submission
Submitted (03-DEC-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs@mail.nih.gov
Tissue procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www.sngc.stanford.edu
Contact: (Dickson, Mark) mcdopaxil@stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAC Plate: 32 Row: f Column: 11.
Location/Qualifiers
1. .1758
/organism="Homo sapiens"
/db_xref="LocusID:51655"
/db_xref="taxon:9606"
/clone="MGC:26290 IMAGE:4797482"
/tissue_type="Brain, hypothalamus"
/clone_lib="NIH_MGC_96"
/lab_host="DH10B"
/note="Vector: pBluescript"
213..1058
/codon_start=1
/product="RAS, dexamethasone-induced 1"
/protein_id="AAH18041.1"
/db_xref="GI:17390076"
/translation="MKLAAMIKKPCPSDSELSIPAKNCYRMVILGSSKVGKTAIVSRF
LTGRDAYITIEDPHRKFSYIRGEVQILDILTSGNHPFPAMRLSLITGDVFILY
FSLNRDSFEVQRRLQIILDTKSCLANKTKENVVDPLVICGNKGDRDFREVDQREI
EQLVGGDPQRCAYFEISAKKNSSLDDQMFRALFMALKLPSEMSPLDHRKVSQYCDVLH
KKALNKILLRAGSGGGGGPDGFAGIVAPFARRPSSHDLMIYREKASAGSQAKDK
RCVIS"

CDS
BASE COUNT 387 a 576 c 501 g 294 t
ORIGIN

Alignment Scores:
Pred. No.: 1.69e-102 Length: 1758
Score: 1721.00 Matches: 330
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.42% Indels: 0
DB: Gaps: 0

US-09-709-103-3F1 (1-332) x BC018041 (1-1758)

Qy 3 ProSerGlyAlaGlyAlaProSerProSerArgAlaGlnProGlnInsProProAla 22
CGCTTCATGCTCCGCGGCCCAAGCCGAGCGGGCCGAGCCAGCAGCCCTTCAGGC 125

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 1740)
Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS

1 (bases 1 to 1740)
Cismowski, M.J., Takesono, A., Ma, C., Lizano, J.S., Xie, X.,
Fuernkranz, H., Lanier, S.M., and Duzic, E.

TITLE

Genetic screens in yeast to identify mammalian nonreceptor
modulators of G-protein signaling

JOURNAL

Nat. Biotechnol. 17 (9), 878-883 (1999)

MEDLINE

9940338

REFERENCE

2 (bases 1 to 1740)

AUTHORS

Cismowski, M.J., Fuernkranz, H., Ma, C., Spruyt, M., Xie, X.,
Lanier, S.M., and Duzic, E.

TITLE

Direct Submision
Submitted (01-JUN-1998) Receptor Pharmacology/Biochemistry, Cadus
Pharmaceutical Corporation, 777 Old Saw Mill River Rd., Tarrytown,
NY 10591, USA

FEATURES

Location/Qualifiers

1..1740

/organism="Homo sapiens"

/db_xref="taxon:9606"

/tissue_type="liver"

/dev_stage="adult"

1..1740

/gene="AGS1"

146..991

/gene="AGS1"

/note="ras-related G protein; similar to Mus musculus
product encoded by Genbank Accession Number AF009246"

/codon_start=1

/product="activator of G protein signaling"

/protein_id="AAD34206.1"

/db_xref="GI:4959038"

/translation="MKLAAMIKMKCSDELSIPAKNCYRWILGSSKVTATVSR
LTCRFDAYPTIEDPKRKYSLRGEVYLDILDSGNHPFAMRLILGDEVITV
FSLDNDSPEEVRLNQIILDTKSLCKNKENNVPLVTCNGKGRDPEVREOREI
EQVYDPPQCAVFEISAKNSLDMFPLPAMAKLPSEMSPLDHRVSVQYCDVH
KKALRNKKILRASSGGGSDPDAGFIVAFARPSSVMSDLMIYRERKASAGSQAQKE
RCVIS"

BASE COUNT

422 a 546 c 483 g 289 t

ORIGIN

Alignment Scores:

pred. No.: 4.72e-102 Length: 1740

Score: 1714.00 Matches: 329

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 99.02% Indels: 0

DB: 9 Gaps: 0

US-09-709-103-3f1 (1-332) x AF069506 (1-1740)

QY 4 SerGlyAlaGlyAlaProSerProSerArgAlaGlnProGluGlnSerProProAlaAla 23
DB 2 AGCGAGCGCGAGCT 61
QY 24 HisProAlaCyHisProSerAspProGlnProLeuSerAlaLeuLeuSerAlaProArg 43
DB 62 CACCCCGGCT 121
QY 44 ProProSerArgProLeuCyProMetLeuLeuAlaMetIleuIleuMetCysPro 63
DB 122 CGGCT 181
QY 64 SerAspSerGluLeuSerIleProAlaAlaSerCysTyrArgMetValIleLeuGlySer 83
DB 182 AGCGAGCGCGAGCT 241
QY 84 SerIleValGlyIleValIleValIleValIleValIleValIleValIleValIleVal 103
DB 242 TCACAGGAGCT 301

QY 104 TyrThrProThrIleGluAspPheHisArgIlePheTyrSerIleArgIleGluValTyr 123
DB 302 TACAGCGCTACCATTCAGAGCGCTTCCACCGCGCGCGCGCGCGCGCGCGCGCGCGCTTAC 361
QY 124 GlnLeuAspIleLeuAspThrSerGlyAsnHisProPheProAlaMetArgArgLeuSer 143
DB 362 CAGCTGCATCTCTCCACAGCGCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTCC 421
QY 144 IleLeuThrGlyAspValPheIleLeuValPheSerLeuAspAsnArgAspSerPheGlu 163
DB 422 ATCTTCACAGAGACCTTTTCATCTCGGTTCAGTTCGACGACCGCGCGCGCGCGCGCT 481
QY 164 GluValGlnArgLeuArgGlnGlnIleLeuAspThrIleSerCysLeuIleAsnIleThr 183
DB 482 GAGGTCAAGCGCTCAGCT 541
QY 184 IysGluAsnValAspValProLeuValIleCysGlyAsnIleGlyAsnIleGlyAsnIle 203
DB 542 AAGGAGAAAGTGAGCGTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTAC 601
QY 204 ArgGluValAspGlnArgGluIleGluGlnLeuValGlyAspAspProGlnArgCysAla 223
DB 602 CGCGAGTGGAGACCGCTTCC 661
QY 224 TyrPheGluIleSerAlaIleValIleValAsnSerSerLeuAspGlnMetPheArgAlaLeuPhe 243
DB 662 TACTTCGAGATCTCGCTTCC 721
QY 244 AlaMetAlaIleValProSerGluMetSerProAspLeuHisArgIleValSerValGln 263
DB 722 GCATGCTTCC 781
QY 264 TyrGlyAspValLeuHisIleValIleValIleValIleValIleValIleValIleVal 283
DB 782 TATCGCACCGTGTGCACAAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTCC 841
QY 284 GlyGlyGlyGlyGlyAspProGlyIleValIleValIleValIleValIleValIleVal 861
DB 842 GCTTCC 901
QY 304 ProSerValHisSerAspLeuMetTyrIleArgGluIleValIleValIleValIleVal 921
DB 902 CCCAGCGTACACAGCTTCC 961
QY 324 LysAspIleGlyAlaArgCysValIleSer 332
DB 962 AAGGAGAAAGCGCGCTGCGTATTCAGC 988

RESULT 3
AF177335
LOCUS 1746 bp mRNA linear PRI 03-OCT-2000
DEFINITION Homo sapiens clone SP1942 unknown mRNA.
ACCESSION AF177335
VERSION AF177335.1 GI:10503968
KEYWORDS
SOURCE
ORGANISM
Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 1746)
Gu, J.R., Wan, D.F., Zhao, X.T., Zhou, X.M., Jiang, H.Q., Zhang, P.P.,
Qin, W.X., Huang, Y., Qiu, X.K., Qian, L.F., He, L.P., Li, H.N., Yu, Y.,
Yu, J. and Han, L.H.
Novel human cDNA clone with function of inhibiting cancer cell
growth
Unpublished
2 (bases 1 to 1746)
Gu, J.R., Wan, D.F., Zhao, X.T., Zhou, X.M., Jiang, H.Q., Zhang, P.P.,
Qin, W.X., Huang, Y., Qiu, X.K., Qian, L.F., He, L.P., Li, H.N., Yu, Y.,
Yu, J. and Han, L.H.
Direct Submision
Submitted (06-AUG-1999) National Laboratory For Oncogenes & Related
Genes, Shanghai Cancer Institute, 25/Ln 2200 Xie Tu Road, Shanghai

US-09-709-103-3F1 (1-332) X AF222979 (1-4990)

QY 272 aLeAlrAspLstysylsleuLeuA rgaLaqlSerGylgylGylGylAsPProGylAs 292
Db 3776 GCTCGGAACAAGAAGCTGCTGGCGCCGCGACGCGCGCGCGCGCGACCCGGCGCA 3835
QY 292 pAlaEPheGylIleValAlapropheAla rgrProSerValHisSerapLeuMetTY 312
Db 3836 CGCCTTGGCATGTGGACACCTTCGGCGCGCGGCCAGCGACGACGACCTCATGTA 3895
QY 312 rIleAlrGluLysAlaSerAlaGylSerGlnAlaYasplGylunrGyrGylIleSe 332
Db 3896 CATCCGCGAAGAGCGCAGCGCGGACGACCGCAAGACCAAGAGCGCTGCGTCATCAG 3955
QY 332 r 332
Db 3956 C 3956

RESULT 5			
AC020558/c			
LOCUS	183334 bp	DNA	linear
DEFINITION	Homo sapiens BAC clone RP11-524F11 from 17,		PRI 09-JAN-2002
ACCESSION	AC020558		complete sequence.
VERSION	AC020558.4		
KEYWORDS	GI:1342397		
SOURCE	HTG.		
ORGANISM	Homo sapiens.		
	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	1 (bases 1 to 183334)		
JOURNAL	Sulston,J.E. and Waterston,R.		
MEDLINE	Toward a complete human genome sequence		
PUBMED	Genome Res. 8 (11), 1097-1108 (1998)		
REFERENCE	9963792		
AUTHORS	9847074		
TITLE	2 (bases 1 to 183334)		
JOURNAL	Mulvaney,E., Maupin,R., LaPlant,Y. and Bielicki,L.		
REFERENCE	The sequence of Homo sapiens BAC clone RP11-524F11		
AUTHORS	3 (bases 1 to 183334)		
TITLE	Waterston,R.H.		
JOURNAL	Direct Submission		
	Submitted (03-JAN-2000) Genome Sequencing Center, Washington		
	University School of Medicine, 4444 Forest Park Parkway, St. Louis,		
	MO 63108, USA		
REFERENCE	4 (bases 1 to 183334)		
AUTHORS	Waterston,R.H.		
TITLE	Direct Submission		
JOURNAL	Submitted (07-MAR-2001) Genome Sequencing Center, Washington		
	University School of Medicine, 4444 Forest Park Parkway, St. Louis,		
	MO 63108, USA		
REFERENCE	5 (bases 1 to 183334)		
AUTHORS	Waterston,R.		
TITLE	Direct Submission		
JOURNAL	Submitted (09-JAN-2002) Department of Genetics, Washington		
	University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA		
COMMENT	On Mar 7, 2001 this sequence version was replaced by		

----- Genome Center
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: <http://genome.wustl.edu/gsc>
 Contact: saplens@wustl.edu
 ----- Summary Statistics -----
 Center project name: H_NH0524F711

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >=

Tue Dec 31 15:27:53 2002

30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RPI1-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P.Y., Zhao, B., Frengen, E., Tateo, M., Cetanese, J.J., and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. *Genomics* 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)

VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is AC073621. Actual start of this clone is at base position 1 of RP11-524F11; actual end is at base position 183334 of RP11-524F11.

The sequence from position 1987 to 2253 was derived from PCR

product of RP11-524F11 BAC DNA.

FEATURES

source

1. .183334
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/map="17"
/clone="RP11-524F11"
/clone_lib="RPGI-11"

misc_feature 1. .339 notes="similar to EST AI672824 (NID:g4852555) w73e06.x1"
misc_feature 339. .887 notes="similar to EST AI968384 (NID:g5765202) wu02e05.x1"
misc_feature 339. .738 notes="similar to EST AI289781 (NID:g391622) qw12a07.x1"
misc_feature 339. .658 notes="similar to EST AW072137 (NID:g6027135) wz9e07.x1"
misc_feature 340. .767 notes="similar to EST AI014398 (NID:g3228779) ou93b12.s1"
misc_feature 340. .686 notes="similar to EST AW237782 (NID:g6570171) xm81e08.x1"
misc_feature 346. .742 notes="similar to EST AI359064 (NID:g4110685) qy25e07.x1"
misc_feature 405. .884 notes="similar to EST AI279211 (NID:g3917445) qm24c09.x1"
misc_feature 782. .1335 notes="similar to EST BF042464 (NID:g10759519)"
misc_feature 822. .1319 notes="similar to EST BE864687 (NID:g10385980)"
misc_feature 834. .2931 notes="CpG island (GC=69.0, o/e=0.80, #CpGs=226)"
misc_feature 1135. .1585 notes="similar to EST BF555582 (NID:g11665312)"
misc_feature 1168. .1583 notes="similar to EST AW028127 (NID:g5886883) wv26c07.x1"
misc_feature 1432. .1583 notes="similar to EST BE390509 (NID:g9335874)"
misc_feature 1440. .1583 notes="similar to EST AW915326 (NID:g8081017)"
misc_feature 1450. .1583 notes="similar to EST AA790463 (NID:g2850583) vw04g12.r1"
misc_feature 1795. .2255 notes="similar to EST BE390509 (NID:g9335874)"
misc_feature 1795. .2201 notes="similar to EST AA790463 (NID:g2850583) vw04g12.r1"
misc_feature 1795. .2153 notes="similar to EST BF555582 (NID:g11665312)"

misc_feature 1795. .2114 notes="similar to EST AW915326 (NID:g8081017)"
misc_feature 1804. .2071 notes="similar to EST AW028127 (NID:g5886883) wv26c07.x1"
misc_feature 1817. .2201 notes="similar to EST BE836084 (NID:g10268462)"
misc_feature 1844. .2201 notes="similar to EST AA672515 (NID:g2644732) vo59e02.r1"
misc_feature 1985. .2285 notes="similar to EST AW463433 (NID:g7033601)"
misc_feature 2538. .3033 notes="similar to EST AA852097 (NID:g2940690)"
misc_feature 2554. .2806 notes="similar to EST BF057610 (NID:g10811506)"
misc_feature 3561. .3924 notes="similar to EST BF195653 (NID:g11082770)"
repeat_region 4027. .4085 notes="similar to EST RI8261 (NID:g771871) yg01a05.x1"
misc_feature 4670. .5117 /rpt_family="MER1_type" notes="similar to EST R42241 (NID:g817108) yf98b09.s1"
misc_feature 4857. .5117 notes="similar to EST T16275 (NID:g518437)"
repeat_region 6155. .6307 /rpt_family="Alu" notes="similar to EST T16275 (NID:g518437)"
repeat_region 6309. .6446 /rpt_family="L2" notes="similar to EST T16275 (NID:g518437)"
repeat_region 6583. .6891 /rpt_family="Alu" notes="similar to EST T16275 (NID:g518437)"
repeat_region 6960. .7153 /rpt_family="L1" notes="similar to EST T16275 (NID:g518437)"
repeat_region 7525. .7838 /rpt_family="Alu" notes="similar to EST T16275 (NID:g518437)"
repeat_region 7851. .8487 /rpt_family="L1" notes="similar to EST T16275 (NID:g518437)"
repeat_region 9000. .9104 /rpt_family="L2" notes="similar to EST T16275 (NID:g518437)"
misc_feature 11463. .11750 notes="similar to EST BF344873 (NID:g11292331)"
misc_feature 11465. .11750 notes="similar to EST AI340185 (NID:g4077112) qo79d07.x1"
misc_feature 11465. .11750 notes="similar to EST AI571390 (NID:g4534764) tn45e12.x1"
misc_feature 11465. .11750 notes="similar to EST AI701138 (NID:g4989038) we10d09.x1"
misc_feature 11465. .11750 notes="similar to EST AI803637 (NID:g5369109) tc43d01.x1"
misc_feature 11465. .11750 notes="similar to EST BE747895 (NID:g10161887)"
misc_feature 11465. .11750 notes="similar to EST BF203429 (NID:g11097015)"
misc_feature 11465. .11697 notes="similar to EST AI869572 (NID:g5543540) w198d03.x1"
misc_feature 11467. .11750 notes="similar to EST AI817466 (NID:g5436545) wk38g01.x1"
misc_feature 11467. .11619 notes="similar to EST AA367144 (NID:g2019483)"
misc_feature 11468. .11750 notes="similar to EST BE218333 (NID:g8905651) hv37c09.x1"
misc_feature 11468. .11750 notes="similar to EST BE549542 (NID:g9791256)"
misc_feature 11468. .11750 notes="similar to EST BE750272 (NID:g10164264)"
misc_feature 11478. .11750 notes="similar to EST BF667874 (NID:g11941769)"
misc_feature 11480. .11717 notes="similar to EST R34866 (NID:g791767) yh86f06.s1"
misc_feature 11483. .11750 notes="similar to EST AA429062 (NID:g2110588) zv49a05.x1"
misc_feature 11541. .11750 notes="similar to EST W38322 (NID:g1320126) zb19g06.r1"
misc_feature 11552. .11750 notes="similar to EST BE910310 (NID:g10406772)"

Tue Dec 31 15:27:53 2002

Percent Similarity:	82.29%	Conservative:	1
Best Local Similarity:	82.04%	Mismatches:	0
Query Match:	95.67%	Indels:	71
DB:	9	Gaps:	1

US-09-709-103-3F1 (1-332) x AF262018 (1-5141)	
QY 3 ProSerGlyAlaGlyAlaProSerProSerArgAlaGlnProGluGlnSerProProLa 22	
Db 2401 CCGAGCGAGCGGAGCGCCCAAGCCCGAGCGCGCCCGCCAGCGCGAGCGCTCCAGCC 2460	
QY 23 AlaHisProAlaCysHisProSerAspProGlnProLeuSerAlaLeuSerAlaPro 42	
Db 2461 GCTACCCCGCGTCCACCCCGAGCCCTCAGCGCGCTCTCGCCCTCTCTCGGCCCGC 2520	
QY 43 ArgProProSerArgProLeuCysProMetIysLeuAlaAlaMetIleIysIysMetCys 62	
Db 2521 CGCCCGCGCTCGCGCCCTCTGCCCAATGAATGCGCGCGATCATCAAGAAGATGTGC 2580	
QY 63 ProSerAspSerGluLeuSerIleProAlaIysAsnCysTyrArgMetValIleLeuGly 82	
Db 2581 CCGAGCGACTCGAGCTGAGTATCCCGCCCAAGAACTGCTATCGCATGGTCCTCCGCGC 2640	
QY 83 SerSerLysValGlyIysThrAlaIleValSerArgPheLeuThrGlyArgPheGluAsp 102	
Db 2641 TCGTCCAGGTGGCAAGCGCCATCGTGTGCGCTTCTTCCACGCGCGCTTCGAGGAC 2700	
QY 103 AlaTyrThrProThrIleGluAspPheHisArgIysPheTyrSerIleArgGlyGluVal 122	
Db 2701 GCCTACAGCGCTACCATCGAGACTTCCACCGCAAGTTCTACTCCATCCGCGCGAGGTC 2760	
QY 123 TyrGlnLeuAspIleLeuAspThrSerGlyAsnHisProPheProAlaMetArgArgLeu 142	
Db 2761 TACCAGCTCGACATCTCGACACGTCGCGCAACCCGCTTCCCGCCCATCGCGCGCTC 2820	
QY 143 SerIleLeuThr----- 146	
Db 2821 TCCATCTCACAGTGAGCGCGCGCGCGCGAGTTCGCGGAGGAGGCGGCGGAAACCT 2880	
QY 146 ----- 146	
Db 2881 CGGCCAGGCGCGCGCGCGCGCGGTCGCGTFCGCGCGCGCGCGAGTAGTGCCTCGCG 2940	
QY 146 ----- 146	
Db 2941 CTTAGAGAGGTAGCGCGCGCGCGCGCTCAAAGTCAGCCGACTTGTCCCTCGCGG 3000	
QY 147 ----- 152	
Db 3001 GCCACCTCACCTTCT 3060	
QY 152 uValPheSerLeuAspAnArgAspSerPheGluGluValGlnArgLeuArgGlnGln 172	
Db 3061 GGTGTTCAGTCGGAACACCGGACTCTCTTCGAGAGGTGCGCGGCTCAAGCAGCAGAT 3120	
QY 172 eLeuAspThrIysSerCysLeuIysAsnLysThrIysGluAsnValAspValProLeuVa 192	
Db 3121 CCTCGACCAAGTCTTGCT 3180	
QY 192 lIleCysGlyAsnLysGlyAspArgAspPheTyrArgGluValAspGlnArgGluIleG 212	
Db 3181 CATCTCGCGCAACAGGGTGACCGGACTTCTACCGCGAGGTGGACCGCGAGATCGA 3240	
QY 212 uGlnLeuValGlyAspAspProGlnArgCysAlaTyrPheGluIleSerAlaLysLysAs 232	
Db 3241 GCAGCTGGTGGCGACGACCCCGCGCTGCGCTACTTTCGAGATCTCGCGCAAGAAGA 3300	
QY 232 nSerSerLeuAspGlnMetPheArgAlaLeuPheAlaMetAlaLysLeuProSerGluMe 252	
Db 3301 CAGCAGCTGGACAGATGTCGCGCGCTCTTCGCGCATGGCCAAAGTCTCCCGCAGGAT 3360	
QY 252 tSerProAspLeuHisArgLysValSerValGlnTyrCysAspValLeuHisLysLysAl 272	

Db 3361 GAGCCGAGACTGACCCCAAGGTTCTGGTCCAGTCTGCGAGCTGCTGCAAGAAGGC 3420	
QY 272 aleuArgAnLysLysLeuLeuArgAlaGlySerGlyGlyGlyGlyGlyGlyGlyGlyGly 292	
Db 3421 GCTGCGACAGAGAGTCTGCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 3480	
QY 292 pAlaPheGlyIleValAlaProPheAlaArgProSerValHisSerAspLeuMetTy 312	
Db 3481 CGCCTTTGGCATCGTGGCACCTTTCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3540	
QY 312 rIleArgGlyLysAlaSerAlaGlySerGlnAlaLysAspLysGluArgCysValIleSe 332	
Db 3541 CATCCGCGAGAGCG 3600	
QY 332 r 332	
Db 3601 C 3601	

RESULT 7	58882 bp	DNA	linear	HTG 19-AUG-2002
AC073621				
LOCUS	Homo sapiens chromosome 17 clone CTD-3073J20 map 17, *** SEQUENCING			
DEFINITION	IN PROGRESS ***, 2 ordered pieces.			
ACCESSION	AC073621			
VERSION	GI:22297517			
KEYWORDS	HTG; HTGS_PHASE2; HTGS_FULLTOP; HTGS_ACTIVEFIN.			
SOURCE	Human.			
ORGANISM	Homo sapiens			
REFERENCE	1 (bases 1 to 58882)			
AUTHORS	Birken, B., Nusbaum, C. and Lander, E.			
TITLE	Homo sapiens chromosome 17, clone CTD-3073J20			
JOURNAL	Unpublished			
REFERENCE	2 (bases 1 to 58882)			
AUTHORS	Birken, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N., Anderson, S., Baldwin, J., Barna, N., Bastien, V., Bedalov, F., Boguslavsky, L., Boukhgalter, B., Brown, A., Burkett, G., Campopiano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Domino, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L., Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J., LaRocque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Lieu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N., McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheeters, R., Meldrim, J., Meneus, L., Mihova, T., Miranda, C., Mlenga, V., Morrow, J., Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Pierre, N., Pisan, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zimmer, A. and Zody, M.			
TITLE	Direct Submission			
JOURNAL	Submitted (27-JUN-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA			
REFERENCE	3 (bases 1 to 58882)			
AUTHORS	Birken, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B., Canarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Fero, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kellis, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., MacLean, C., Macdonald, P., Major, J., Mathews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Morrow, J., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Pisan, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zimmer, A. and Zody, M.			

Qy 147 -----

Tue Dec 31 15:27:53 2002

```

661 GAGATGAGCCAGCTGACCGCAGGCTCGGTGTCAGTACTGCGCAGCTGTCGACAAAG 720
QY 9 ProSerProSerArgAlaGlnProGlu-GlnSerProProAlaAlaHisProAlaCysHi 28
271 LysAlaLeuArgAsnLysLysLeuLeuArgAlaGlySerGlyGlyGlyGlyGlyGlyGly 75
Db 27 CCAGACAGCTCGCGCAGTCCGAAGCAAGCTCTCTCCGTCACTCCCG----- 75
281 AAGCGCTGCGGAACAAGAGCTGCTCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 780
QY 28 sProSerAspProGlnProLeuSerAlaLeuLeuSerAlaProArgProProSerArgPr 48
291 GlyAspAlaPheGlyLeuAlaProPheAlaArgProSerValHisSerAspLeu 310
Db 291 GGCACGCTTTGGCATCTGGACCTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 840
QY 311 MetTyrLeuArgGluLysAlaSerAlaGlySerGlnAlaLysAspLysGluArgCysVal 330
Db 841 ATGTACATCCGAGAGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGCGTC 900
QY 331 IleSer 332
Db 901 ATCAGC 906

RESULT 10
LOCUS AF009246 1623 bp mRNA linear ROD 13-PEB-1998
DEFINITION Mus musculus ras-related protein (DEXRAS1) mRNA, complete cds.
ACCESSION AF009246
VERSION AF009246.1 GI:2253712
KEYWORDS
SOURCE Mus musculus.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1623)
Kempainen,R.J. and Behrend,E.N.
Dexamethasone rapidly induces a novel ras superfamily
member-related gene in Atci-20 Cells
J. Biol. Chem. 273 (6), 3129-3131 (1998)
98123070
MEDLINE
PUBMED 9452419
REFERENCE
2 (bases 1 to 1623)
Kempainen,R.J. and Behrend,E.N.
Direct Submission
AUTHORS
TITLE Submitted (18-JUN-1997) Physiology & Pharmacology, Auburn
JOURNAL University College of Veterinary Medicine, 213 Greene Hall, Auburn,
AL 36849, USA
FEATURES
source
1..1623
/organism="Mus musculus"
/db_xref="taxon:10090"
/cell_line="Corticotrope tumor cell line Att-20"
1..1623
/gene="DEXRAS1"
142..984
/gene="DEXRAS1"
/notes="induced by dexamethasone"
/codon_start=1
/product="ras-related protein"
/protein_id="AAC53538.1"
/db_xref="GI:2253713"
/translatability="MKLAAMIKMKWCPSPDSELSIPAKNCYRMVILGSSKVGKTAIVSRF
LTFGRFAYPTIEDHFKFYSIRGEVYQILDITSGNHPFFAMRRLSILITGDFVILV
FSLNRDSFEVRLKQIILDTKSLCKNTKENVDVPLVTCGGRDRFYRVEOREI
EQLVGDPORCAYFEISAKNSSLDMFRALFAMAKLPSEMSDHLRKVSQVQICVLH
KVALNKLLRAGSGGGDHGDFGILAPFARRPSVHSDLMVIREKTSVGSQAKOKER
CVIS"
BASE COUNT 412 a 437 c 417 g 357 t
ORIGIN
Alignment Scores: 3.87e-86 Length: 1623
Pred. No.: 1467.00 Matches: 293
Score: 92.31% Conservative: 7
Percent Similarity: 90.15% Mismatches: 18
Best Local Similarity: 84.75% Indels: 7
Query Match: 10 Gaps: 2
DB:

```

```

US-09-709-103-3F1 (1-332) x AF009246 (1-1623)
QY 9 ProSerProSerArgAlaGlnProGlu-GlnSerProProAlaAlaHisProAlaCysHi 28
Db 27 CCAGACAGCTCGCGCAGTCCGAAGCAAGCTCTCTCCGTCACTCCCG----- 75
QY 28 sProSerAspProGlnProLeuSerAlaLeuLeuSerAlaProArgProProSerArgPr 48
Db 76 ---AGTTCTCTGCGAGCCCTCTCTTACCTTCTCTCGGCGCAGCATCCGCCCTGGGGCC 131
QY 48 oLeuCysProMetLysLeuAlaAlaMetLysLysMetCysProSerAspSerGluLe 68
Db 132 CCTTACCCCATGAACTGCGCGCGATGATCAAGAAGATGTGCCCAAGCGACTCTGAAC 191
QY 68 uSerLleProAlaLysAsnCysTyrArgMetValLleLeuGlySerSerLysValGlyLy 88
Db 192 GAGTATCCCGGCAAGAACTGTCTACAGGATGGTGTATCTCTCGGCTCATTCACCAAGTGGCAA 251
QY 88 sThrAlaLleValSerArgPheLeuThrGlyArgPheGluAspAlaTyrThrProThrIl 108
Db 252 GACGGCCATTTGTTCGGCGCTTCTCTACGGGCGGTTTCGAGGATGCTTACACCCCTACCAT 311
QY 108 eLluAspPheHisArgLysPheTyrSerLleArgGlyGluValTyrGlnLeuAspLleLe 128
Db 312 CGAGGACTTCCACCGAAAGTTTACTCGATCCCGCGGGAAGTCTACAGTTGGACATCT 371
QY 128 uAspThrSerGlyAsnHisProPheProAlaMetArgArgLeuSerLleLeuThrGlyAs 148
Db 372 GGACACATCCCGGCAATCATTCGTTTCCCGCATCGCGCGCTCTCTATCTCACAGGAGA 431
QY 148 pValPheLleLeuValPheSerLeuAspAsnArgAspSerPheGluGluValGlnArgLe 168
Db 432 CGTTTTCATCTTGTGTGTTTACGCTTAGACAACCGCGACTCATTCGAGAGGTGGCAAGGCT 491
QY 168 uArgGlnGlnLleLeuAspThrLysSerCysLeuLysAsnLysThrLysGluAsnValAs 188
Db 492 CAAACAGCAGATCTTAGACACCAAGTCTCTCTCAGAACAAAACCAAGAGATGTGGA 551
QY 188 pValProLeuValLleCysGlyAsnLysGlyAspArgAspPheTyrArgGluValAspGl 208
Db 552 CGTGCCTGCTGTTCAITTCGGGTAAACAAAGGGGACCGGAGCTTCTACCGGGAGTAGACA 611
QY 208 nArgGluLleGluGlnLeuValGlyAspAspProGlnArgCysAlaTyrPheGluLleSe 228
Db 612 GCGGAGATTGACAGCTGTTGGGTGACGACCTCAGCGTTGTCCTACTTTCGAGATCTC 671
QY 228 rAlaLysLysAsnSerSerLeuAspGlnMetPheArgAlaLeuPheAlaMetAlaLysLe 248
Db 672 AGCCAGAGAAACAGCAGCTTGGACCAAGATGTTCCGTGCGCTCTTTGGCATGGCCAGCT 731
QY 248 uProSerGluMetSerProAspLysHisArgLysValSerValGlnTyrCysAspValLe 268
Db 732 GCCTAGCGAGATGAGCCCGGCTTGCACCGCAAGAGTATCTGTGCAGTACTGCGAGCTACT 791
QY 268 uHisLysLysAlaLeuArgAsnLysLysLeuLeuArgAlaGlySerGlyGlyGlyGly 288
Db 792 GCACAAAGAGCTCTGAGGAAACAAAGAGCTTCTCGTGGCGGGCAGC---GGAGGGCGGGG 848
QY 288 yAspProGlyAspAlaPheGlyLleValAlaProPheAlaArgArgProSerValHisSe 308
Db 849 CGACACCGGCGATGCCTTTGGCATCTTGGCGCCCTTTTGTTCGAGACCCAGCGGTGCACAG 908
QY 308 rAspLeuMetTyrLleArgGluLysAlaSerAlaGlySerGlnAlaLysAspLysGluAr 328
Db 909 CGACCTCATGTACATTCGTGAAAACACCAAGTGTTCGCGAGCGCAGGCTAAGGACAGGCG 968
QY 328 gCysValLleSer 332
Db 969 CTGTGTCTCATCAGT 981
RESULT 11
BC034166 1612 bp mRNA linear ROD 07-AUG-2002
LOCUS

```

DEFINITION Mus musculus, RAS, dexamethasone-induced 1, clone MGC:36188
IMAGE:498312, mRNA, complete cds.
ACCESSION BC034166
VERSION BC034166.1 GI:21706874
KEYWORDS MGC.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1612)
Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (02-JUL-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mdickpaxil@stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRAX Plate: 58 Row: 9 Column: 5
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 6677672.
Location/Qualifiers
1. 1612
/organism="Mus musculus"
/db_xref="taxon:10090"
/map="FVB/N"
/clone="MGC:36188 IMAGE:498312"
/tissue_type="Colon normal, 5 month old male mouse."
/clone_lib="NCI CGAP_C024"
/lab_host="DH10B"
/note="Vector: PCMV-SPORT6"
108..950
/codon_start=1
/product="RAS, dexamethasone-induced 1"
/protein_id="AAH34166.1"
/db_xref="GI:21706875"
/db_xref="LocusID:19416"
/translation="MKLAMIRKMCPSSELSIPAKNCYRWVILGSSKVGKTAIVSRF LTRFEDAVPTIEDPHRKFSIRGVQOLDLDSGNHPPRMRSLITGDFVILV FSLDNDSFEVQRLKQIILDTKSCKKTKENUNPVITGNGGDFIFREYQDEIV EQLVDDPORCAVEFISAKNSLDPMRALFPAKKLPSSESPULHKKVAVQYQDEIV KKAIRKKLLRAGSGGGGHDAPGIIAPFAPRPSVSDLMYIREKTSVGSQAKDER CVTS"
BASE COUNT 417 a 424 c 420 g 351 t
ORIGIN
Alignment Scores:
Pred. No.: 2,86e-85 Length: 1612
Score: 1453.50 Matches: 290
Percent Similarity: 91.67% Conservative: 7
Best Local Similarity: 89.51% Mismatches: 16
Query Match: 83.97% Indels: 11
DB: 10 Gaps: 3
US-09-709-103-3f1 (1-332) x BC034166 (1-1612)
QY 11 ProSerArgAlaGlnProGluInSerProProAlaAlaHisProAlaCySHsProSer 30
Db 3 CCAGCTCGCGCAGTCCCGAAGCAAGCTCTTCCTGCTACT----- 41

QY 31 AspProGlnProLeuSerAla-----LeuLeuSerAlaProArgProProSerArgPro 48
Db 42 ---CGGAGTTCTTGACGACCTCTACCTCTCTCGGCGACGACCTCGGGGCC 98
QY 49 LeuCyProMetCysLeuAlaAlaMetIleLysLysMetCysProSerAspSergIleu 68
Db 99 CTCACCAATGAAGTGGCCCGGATGATCAAGAGATGTGCCCAACGACTCTGAATG 158
QY 69 SerIleProAlaLysAsnCySTyrArgMetValIleuGlySerSerIleValGlyLys 88
Db 159 AGTATCCGCGCCAAAGTCTGACAGATGTCTCTCGGCTCATCCAAAGTGGCAAG 218
QY 89 ThrAlaIleValSerArgPheLeuThrGlyArgPheLysAlaTyrThrProthrlle 108
Db 219 ACGGCATTGTGTGCGCTTCTCTCAAGCGGCTTTCAGAGATCTTCAACCCCTACATC 278
QY 109 GluAspPheHisArgLysPheTyrSerIleArgGlyGluValTyrGlnLeuAspIleu 128
Db 279 GAGGACTTCCACCAAGTTTACTCGATCCGGGGGAAGTCTACCAAGTTGACATACG 338
QY 129 AspThrSerGlyAsnHisProPheProAlaMetArgArgLeuSerIleLeuThrGlyAsp 148
Db 339 GACCATCCGCGCATCATCCGTTTCCGCGCATGCGGCTCTATCTCTACAGAGAC 398
QY 149 ValPheIleLeuValPheSerLeuAspAsnArgAspSerPheGluGluValGlnArgLeu 168
Db 399 GTTTCATTCTGCTGTTCACTTACGACCAACCGGACTCTTGAAGAGTGCAAAGGCTC 458
QY 169 ArgGlnIleLeuAspThrLysSerCysLeuLysAsnLysThrLysGluAsnValAsp 188
Db 459 AAACAGAGATCTTCAAGCACCAAGTCTGTCTCAAGAACCAAGAGATGTGGAC 518
QY 189 ValProLeuValIleCysGlyAsnLysGlyAspArgAspPheTyrArgGluValAspGln 208
Db 519 GTGCGCTGCTGCTTGGCGGTAAACAAAGGAGCCGGACTCTTACCGGAAAGTAGACAG 578
QY 209 ArgGlnIleLeuGlnLeuValGlyAspAspProGlnArgCysAlaTyrPheGluIleSer 228
Db 579 CCGGAGATTGAGCAGCTGTGGGTGACACCTCAGCGTGTGCTTCTCGAGATCTCA 638
QY 229 AlaLysLysAsnSerSerLeuAspGlnMetPheArgAlaLeuPheAlaMetAlaLysLeu 248
Db 639 GCCAAGAAAGACAGCAGCTTGACCAAGATGCTTCCGCTCTTGTGCATGGCCAAAGCTG 698
QY 249 ProSerGluMetSerProAspLeuHisArgLysValSerValGlnTyrCysAspValLeu 268
Db 699 CTTAGGAGATGAGCCCGGACTTGCACCGCAAGTATCTGTGAGTACTGCAAGCTACTG 758
QY 269 HisLysLysAlaLeuArgAsnLysLysLeuLeuArgAlaGlySerGlyGlyGlyGly 288
Db 759 CACAAGAAAGCTCTGAGGAAACAAGAACTTCTGCGGCGGCGC---GGAAGCGGGGCG 815
QY 289 AspProGlyAspAlaPheGlyIleValAlaProPheAlaArgArgProSerValHisSer 308
Db 816 GACCAAGCGGAGTCTTGGCATCTTGGCGGCTTGTCTCGGAGCCAGCTGACAGCC 875
QY 309 AspLeuMetTyrIleArgGlyLysAlaSerAlaGlySerGlnAlaLysAspLysGlnArg 328
Db 876 GACCTCATGTACATTCGTGAAAAACCAAGTGTGGGCGCAGGCTTAAGACAAAGAGCGCG 935
QY 329 CysValIleSer 332
Db 936 TGTGTCACTCACT 947
RESULT 12
AF498923
LOCUS AF498923
DEFINITION Homo sapiens activator of G protein signaling (RASD1) mRNA, PRI 01-MAY-2002
ACCESSION AF498923
VERSION AF498923.1 GI:20379021
KEYWORDS Homo sapiens.
SOURCE

Tue Dec 31 15:27:53 2002

```

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 846)
AUTHORS Puhl H.L., Ikeda S.R. and Aronstam, R.S.
TITLE Direct Submission
JOURNAL Submitted (05-APR-2002) cDNA Resource Center, Guthrie Research
Institute, One Guthrie Square, Sayre, PA 18840, USA
FEATURES
    source
        1..846
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /tissue_type="brain"
    gene
        1..846
            /gene="RASD1"
            /note="AGS1"
    CDS
        1..846
            /gene="RASD1"
            /codon_start=1
            /product="activator of G protein signaling"
            /protein_id="AAM21071.1"
            /db_xref="GI:20379022"
            /translation="MKLAAMIKMCPDSELSIPAKNCYRMVILGSSKVGKTAIVSRF
            LTGRFEDATPTIEDFHRKFYSIRGEVQILDITSGNHPFPFAMRRLSILITGDFVILV
            FSLNDRDSFEVORLROQILDITKSLNKTENVDPLVICGNKGRDFYREVDOREI
            EQLVDDPQRCAYFEISAKNSLDDQMFALFAMAKLPSEMSPLDHRKVSQVCDVLH
            KKALRNKLLRAGSGGGGPDGAFGIVAFARRPSVHSDLMYIREKASAGSQAKDKE
            RCVIS"
BASE COUNT 176 a 288 c 252 g 130 t
ORIGIN
Alignment Scores: 2.79e-85 Length: 846
Pred. No.: 1449.00 Matches: 281
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 83.71% Gaps: 0
DB: 9
US-09-709-103-3f1 (1-332) x AF498923 (1-846)
QY 52 MetLysLeuAlaMetLleLysLysMetCysProSerAspSerGluLeuSerIlePro 71
Db 1 ATGAACCTGGCGGAGTATCAAGAGATGTCCCGGACGACCTCGAGGTGAGTATCCCG 60
QY 72 AlaLysAsnCysTyrArgMetValIleuGlySerSerLysValGlyLysThrAlaIle 91
Db 61 GCCAAGAACTGCTATCGCATGCTCTCTCGCTCTCGTCCAGGTGGGCAAGACGGCATC 120
QY 92 ValSerArgPheLeuThrGlyArgPheGluAspAlaTyrThrProThrIleGluAspPhe 111
Db 121 GTGTGCGGCTTCTCACCGCGGCTTCGAGACGCTACAGCTACATCCATCGAGGACTTC 180
QY 112 HisArgLysPheTyrSerIleArgGlyGluValTyrGlnLeuAspIleLeuAspThrSer 131
Db 181 CACCGCAAGTTTCTTACTCCATCCGCGGAGGTCTACACAGCTCGACATCTCTCACAGCTCC 240
QY 132 GlyAsnHisProPheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIle 151
Db 241 GGCAACACCCGTTCCCGCCATCGGCGCTCTCCATCCTCACAGGAGAGCTTTTCATC 300
QY 152 LeuValPheSerLeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGlnGln 171
Db 301 CTGGTGTTCAGTCTGACACACCGGACTCTTCGAGGAGGTGACGGCTCAGGCGAGCAG 360
QY 172 IleLeuAspThrLysSerCysLeuLysAsnLysThrLysGluAsnValAspValProLeu 191
Db 361 ATCTTCGACACCAAGTCTTCCTCAAGAACAAACCAAGGAGAACCTGAGCTGCCCTG 420
QY 192 ValIleCysGlyAsnLysGlyAspArgAspPheTyrArgGluValAspGlnArgGluIle 211
Db 421 GTCATCTCGGCAACAAAGGTGACCGGACTTCTTACCGGAGGTGACCGAGGATC 480

```

AF153192 1187 bp mRNA linear PRI 02-JUN-1999
Homo sapiens ras-related protein mRNA, complete cds.

AF153192
AF153192
AF153192.1 GI:4960166

KEYWORDS
SOURCE
ORGANISM

Homo sapiens
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1187)
Kempainen, R.J.
Identification of human pituitary Dextral

Unpublished
2 (bases 1 to 1187)

REFERENCE
AUTHORS
JOURNAL
TITLE
JOURNAL

Direct Submission
Submitted (20-MAY-1999) Anatomy & Physiology, Auburn University,
College of Veterinary Medicine, Auburn, AL 36849, USA
Location/Qualifiers
1..1187

FEATURES
source

/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="2"
/map="2q32"

/tissue_type="pituitary gland"

CDS
1..846
/notes="Dextral; similar to Mus musculus ras-related
protein encoded by GenBank Accession Number AF009246"

/codon_start=1
/product="ras-related protein"

/protein_id="AAB3421.1"

/db_xref="GI:4960167"
/translations="MKLAAMIKMCPDSELSIPAKNCYRMVILGSSKVGKTAIVSRF
LTGRFEDATPTIEDFHRKFYSIRGEVQILDITSGNHPFPFAMRRLSILITGDFVILV
FSLNDRDSFEVORLROQILDITKSLNKTENVDPLVICGNKGRDFYREVDOREI
EQLVDDPQRCAYFEISAKNSLDDQMFALFAMAKLPSEMSPLDHRKVSQVCDVLH
KKALRNKLLRAGSGGGGPDGAFGIVAFARRPSVHSDLMYIREKASAGSQAKDKE
RCVIS"

BASE COUNT 249 a 396 c 362 g 180 t
ORIGIN

Alignment Scores: 4.01e-85 Length: 1187
Pred. No.:

109	Qy	109	GluAspPheHisArgIysPheTyrSerIleArgGlyGluValTyrGlnLeuAspIleLeu	128
289	Db	289	GAAGACATTCACCGAAGATTTCATCTCGATCCGCGCGAAGTCTACCAAGTTGGACATACCTG	348
129	Qy	129	AspThrSerGlyAsnHisProPheProAlaMetArgArgLeuSerIleLeuThrGlyAsp	148
349	Db	349	GACACATTCGGCAATCATCGTTTCCCGCATCGCGCGCTCTATCTCTCACAGGAGAC	408
149	Qy	149	ValPheIleLeuValPheSerLeuAspAsnArgAspSerPheGluGluValGlnArgLeu	168
409	Db	409	GTATTTCAATCTCTGGTGTTCAGCTTAGACACACCGCGACTCTCTCGAGGAGGTGCAAGGCTC	468
169	Qy	169	ArgGlnGlnIleLeuAspThrLysSerCysLeuLysAsnIleThrLysGluAsnValAsp	188
469	Db	469	AACACAGCAGATCCTAGACACCACTGCTGTCTCAAGAACAAACCAAGAGATGTGGAC	528
189	Qy	189	ValProLeuValIleCysGlyAsnLysGlyAspAspAspPheTyrArgGluValAspGln	208
529	Db	529	GTGCCGCTGGTCAATTTGGGTAAACAAAGGGGACCGGGACTTCTACCGCAAGTGGAGCAG	588
209	Qy	209	ArgGluIleGlnLeuValGlyAspAspProGlnArgCysAlaTyrPheGluIleSer	228
589	Db	589	CGGAGATTTGACGAGCTGTGTGGCGATGACCCCTCAGCGTTGTGCCCTACTTCGAGATCTCG	648
229	Qy	229	AlaLysLysAsnSerSerLeuAspGlnMetPheArgAlaLeuPheAlaMetAlaLysLeu	248
649	Db	649	GCCAAAGAAGATAGCAGCTGGACCAAGATGTTCCGTGGCTCTTTGCCATGGCCAGCTG	708
249	Qy	249	ProSerGluMetSerProAspLeuHisArgLysValSerValGlnTyrCysAspValLeu	268
709	Db	709	CTTAGCAGATGAGCCCTGACTTGCACCGCAAGGTGTCTGTGCAGTACTGTGACGTGCTG	768
269	Qy	269	HisLysLysAlaLeuArgAsnLysLysLeuLeuArgAlaGlySerGlyGlyGlyGly	288
769	Db	769	CACAAAAGGCTCTGAGGACCAAGAGCTTCTGCTGCGGGCAGC--GGAGGTGGGGGC	828
289	Qy	289	AspProGlyAspAlaPheGlyIleValAlaProPheAlaArgProSerValHisSer	308
826	Db	826	GACCACGGAGATGCTTTGGCATCTTTGGCGCCCTTTGTCTCGACACCTAGCGTGCATAGC	885
309	Qy	309	AspLeuMetTyrIleArgGluLysAlaSerAlaGlySerGlnAlaLysAspLysGluArg	328
886	Db	886	GACCTCATGTACATTCGTGTGAGAAAACCGTGTGACGACCGAGCTTAGGACAAAGGAGCGC	945
329	Qy	329	CysValIleSer	332
946	Db	946	TGTGTCAATCAGT	957
RESULT 15				
AL603710/c				
LOCUS				
DEFINITION				
Mouse DNA sequence from clone RP23-247B13 on chromosome 11,				
complete sequence.				
ACCESSION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
house mouse.				
Mus musculus				
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
REFERENCE				
1 Almeida J.				
Direct Submission				
Submitted (04-APR-2002) Wellcome Trust Sanger Institute, Hinxton,				
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:				
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk				
On Apr 7, 2002 this sequence version replaced gi:18070899.				
During sequence assembly data is compared from overlapping clones.				
Where differences are found these are annotated as variations				
together with a note of the overlapping clone name. Note that the				
variation annotation may not be found in the sequence submission				
corresponding to the overlapping clone, as we submit sequences with				
only a small overlap as described above.				
COMMENT				

Db 122039 TGCAGCTCTGCTGCGCAGCTGCTCCTCCTTCCACTGCTTCCCTTATAGGAGCGT 121980
QY 149 lPhe1leLeuValPheSerLeuAspAsnArgAspSerPheGluGluValGlnArgLeuAr 169
Db 121979 TTTCAATTCTGGTGTCTTACGCTTGAACACCGGACCTCATTCGAAAGAGGTGCMAAGGCTCA 121920
QY 169 gGlnGln1leLeuAspThrLysSerCysLeuLysAsnLysThrLysGluValAspVal 189
Db 121919 ACAGCGATCTTACAGACACCAAGTCTGTCTCAAGAACMAACMAAGAGATGTGACGT 121860
QY 189 lProLeuVal1leCysGlyAsnLysGlyAspArgAspPheTyrArgGluValAspGlnAr 209
Db 121859 GCGGCTGGTCACTTCCGCTTAACAAGGGGACCGGACCTTCAACGGGAGTAGAGCAGCG 121800
QY 209 gGlu1leGluGluLeuValGlyAspAspProGlnArgCysAlaTyrPheGlu1leSerAl 229
Db 121799 GGAGATTGAGCAGCTGGTGGGTGACACCTCAAGCTTGTGCTTCTCGAGATCTCAGC 121740
QY 229 aLysLysAsnSerSerLeuAspGlnMetPheArgAlaLeuPheAlaMetAlaLysLeuPr 249
Db 121739 CAAGAAAGACAGCAGCTTGGACAGATGTTCCGTGGCTCTTGGCCATGCGCAAGCTGCC 121680
QY 249 cSerGluMetSerProAspLeuHisArgLysValSerValGlnTyrCysAspValLeuHi 269
Db 121679 TAGCGAGATGAGCCCGACTTGGACCGCAAGGTATCTGTGCAGTACTGCGAGTACTGCA 121620
QY 269 sLysLysAlaLeuArgAsnLysLysLeuLeuArgAlaGlySerGlyGlyGlyGlyAs 289
Db 121619 CAAGAAAGCTCTTGAAGAACMAAGCTTCTGCGTGGCGGCAAC---GGAGCGGGGGCGCA 121563
QY 289 pProGlyAspAlaPheGly1leValAlaProPheAlaArgArgProSerValHisSerAs 309
Db 121562 CCAAGGCGATGCTTGGATCTTGGCGCCCTTGTGCTGCAAGCCAGCGTGCACAGCA 121503
QY 309 PleuMetTyr1leArgGluLysAlaSerAlaGlySerGlnAlaLysAspLysGluArgCy 329
Db 121502 CTTCACTGACATTCGTAAGAAAAACAGTGTGCGGAGCCAGGCTAAGGACAAAGAGCGCTG 121443
QY 329 sVal1leSer 332
Db 121442 TGTCACTCAGT 121433

Search completed: December 30, 2002, 19:08:33
Job time : 3287 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: December 30, 2002, 16:29:50 ; Search time 87 Seconds

(without alignments)
1316.766 Million cell updates/sec

Title: US-09-709-103-1f4

Perfect score: 1411
Sequence: 1 LAADALLVIGLAAGAGLLA.....RDTOLVARAHLDHRCQFH 282

Scoring table:

BLOSUM62
Xgapod 10.0 , Xgapext 0.5
Ygapod 10.0 , Ygapext 0.5
Fgapod 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 363474 segs, 203117208 residues

Total number of hits satisfying chosen parameters: 726948

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+g2n.model -DEV=x1p
-Q=/cgn2_1/USPTO.spool/SUPLIVANI14/runat_30122002_144047_21556/appd_query.fasta_1.455
-DB=Published Applications NA -OPMT=fastap -SUFFIX=g2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=D1osum62
-TRANS=humano.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=SUPLIVANI14@CGN_1_1_36 @runat_30122002_144047_21556
-DEPU=6 -ICPU=3 -NO_XLPHY -NO_MMAP -LARGEQUERY -NEG SCORES=0 -MATT -LONGLOG
-DEVTIMOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	575.5	40.8	3082	10	US-09-778-963A-1
C 2	397.5	28.1	405	10	US-09-960-352-10273
C 3	371.5	26.3	11221	10	US-09-778-963A-3
C 4	219	15.5	368	10	US-09-864-761-21643

C 5	164.5	11.7	452	10	US-09-960-352-4253	Sequence 4253, Ap
C 6	160.5	11.4	454	10	US-09-960-352-5830	Sequence 5830, Ap
C 7	159.5	11.3	2322	10	US-09-476-242-20	Sequence 20, Appl
C 8	155	11.0	2322	10	US-09-476-242-18	Sequence 18, Appl
C 9	155	11.0	2324	10	US-09-476-242-7	Sequence 7, Appl
C 10	154	10.9	1224	10	US-09-972-529-1	Sequence 1, Appl
C 11	154	10.9	2319	10	US-09-788-654A-1	Sequence 1, Appl
C 12	153	10.8	2322	10	US-09-476-242-19	Sequence 19, Appl
C 13	146.5	10.4	2010	12	US-10-032-717-9	Sequence 9, Appl
C 14	146.5	10.4	4074	10	US-09-815-242-7899	Sequence 7899, Ap
C 15	143	10.1	612	10	US-09-476-242-12	Sequence 12, Appl
C 16	143	10.1	2541	10	US-08-841-635A-32	Sequence 32, Appl
C 17	142	10.1	1894	7	US-09-934-060A-5	Sequence 5, Appl
C 18	139	9.9	1668	9	US-09-476-242-26	Sequence 26, Appl
C 19	139	9.9	2352	10	US-09-476-242-9	Sequence 9, Appl
C 20	139	9.9	2541	10	US-09-476-242-10	Sequence 10, Appl
C 21	139	9.9	2541	10	US-09-476-242-11	Sequence 11, Appl
C 22	139	9.9	2541	9	US-09-860-846-3	Sequence 9, Appl
C 23	137.5	9.7	2159	9	US-09-861-289-9	Sequence 9, Appl
C 24	137	9.7	1458	10	US-09-815-242-7835	Sequence 7835, Ap
C 25	137	9.7	1629	10	US-09-815-242-4085	Sequence 4085, Ap
C 26	137	9.7	1707	10	US-09-476-242-21	Sequence 21, Appl
C 27	137	9.7	2310	10	US-09-861-289-3	Sequence 3, Appl
C 28	137	9.7	2535	10	US-09-861-289-3	Sequence 3, Appl
C 29	137	9.7	13613	9	US-10-025-380-1058	Sequence 1058, Ap
C 30	137	9.7	13613	10	US-09-922-217-1058	Sequence 1058, Ap
C 31	136.5	9.7	15720	10	US-09-833-263-1058	Sequence 1058, Ap
C 32	136.5	9.7	15720	10	US-10-104-484-1	Sequence 1, Appl
C 33	136.5	9.7	570	12	US-09-476-242-22	Sequence 22, Appl
C 34	136.5	9.7	570	12	US-10-124-800-17	Sequence 17, Appl
C 35	136	9.6	2258	10	US-09-476-242-22	Sequence 22, Appl
C 36	136	9.6	2133	9	US-10-124-800-17	Sequence 17, Appl
C 37	136	9.6	2133	9	US-09-476-242-22	Sequence 22, Appl
C 38	135.5	9.6	8730	10	US-09-476-242-24	Sequence 24, Appl
C 39	135.5	9.6	2298	10	US-09-476-242-15	Sequence 15, Appl
C 40	135	9.6	2298	10	US-09-476-242-14	Sequence 14, Appl
C 41	135	9.6	2523	10	US-10-124-800-5	Sequence 5, Appl
C 42	135	9.6	2523	10		
C 43	135	9.6	4509	9		
C 44	135	9.6				
C 45	135	9.6				

ALIGNMENTS

RESULT 1
US-09-778-963A-1/c
Sequence 1, Application US/09778963A
Patent No. US20020115172A1
GENERAL INFORMATION:
APPLICANT: NEELAM, Beena et al
TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS, AND USES THEREOF
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE PROTEINS, AND USES THEREOF
FILE REFERENCE: CLO01112
CURRENT APPLICATION NUMBER: US/09/778, 963A
CURRENT FILING DATE: 2001-02-08
NUMBER OF SEQ ID NOS: 5
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 3082
TYPE: DNA
ORGANISM: Homo sapiens
US-09-778-963A-1

Alignment Scores:

Pred. No.: 3.05e-44
Score: 575.50
Percent Similarity: 61.23%
Best Local Similarity: 52.54%
Query Match: 40.79%
DB: 10
Length: 3082
Matches: 145
Conservative: 24
Mismatch: 94
Indels: 13
Gaps: 5

Tue Dec 31 15:27:52 2002

```

US-09-709-103-1F4 (1-282) x US-09-778-963A-1 (1-3082)
Qy 4 AspAlaAlaLeuValLeu---GlyLeuAlaAlaGlyAlaGlyLeuLeuAlaAspVal 22
Db 1005 GATGTCACCTTGCCTTCCACGGCGCTGCGCTTCCGAGGACCTTGGCCCTTGATGTA 946
Qy 23 HisGluValAlaValTyrAlaGlyProAlaArgGluGlyCysHisAspAlaLysGlyVal 42
Db 945 CTTGAGGTCACTGTGACGCTGGCGCGCGGGCGGCGAGACCATTCATAGGCGTC 886
Qy 43 AlaArgValAlaAlaAlaAlaAlaAlaGlyProGlnGlnLeuValProGlnArg 62
Db 895 ---CATCTCTTCCGCGCGCATGCAGAGGCGCTGGG----- 850
Qy 63 LeuLeuValGlnHisValAlaValLeuHisArgAspLeuAlaValGlnValTyrAlaHis 82
Db 849 -----GTGGAAGGCGTCACTGACGATCTGCGGATCTTGGCATGCGGGCGGCTCAT 796
Qy 83 LeuAlaGlyGlnLeuGlyHisGlyGluGluArgAlaGluHisLeuValGlnAlaVal 102
Db 795 CTCGTGTGGCAGCTTGGCCATGCTGAACAGCAGCAGTAGAACATCTCGTCCACGTTGGTGT 736
Qy 103 LeuLeuGlyArgAspLeuValGlyAlaAlaLeuGlyValValAlaHisGlnLeuLeu 122
Db 735 CTTCTTGGCGGACACTCGAAGTAGCGGAGTCTCT---GTGCGCCGACACGACGCTC 679
Qy 123 AspLeuAlaLeuValHisLeuAlaValGluValAla---ValThrLeuValAlaAlaAsp 141
Db 678 GGCCTCGGTGGGACCTGGCGGACAGCTGCGCGGTGCTCTTGTGTCACAGAT 619
Qy 142 AspGlnGlyHisValHisValLeuLeuGlyPheValLeuGluAlaArgLeuGlyValGlu 161
Db 618 GACCATGGGCGACTCGCGCGCTCTTGTGCTCTTCTTCCAGGAGGACTTGAACCTCCAG 559
Qy 162 AspLeuLeuProGlnProLeuHisLeuLeuGlyValAlaValAlaValGlnThrGluHis 181
Db 558 GATCTGCTCTGAAGCGCTTGAACCTCATCGAAGGACTCCCGGTATCCAGCTGAACAC 499
Qy 182 GlnAspGluAsnValSerCysGluAspGlyGluAlaProHisGlyGlyGluArgValVal 201
Db 498 CAGGATGAAGACATCCCTGTGAGGATGACAGCCTGCGCATGGCGGGAAGGGTGT 439
Qy 202 AlaGlyArgValGluAspValGluLeuValAspLeuAlaAlaAspGlyValGluLeuAla 221
Db 438 GCCAGAGGTATCCAGGATGTCAGCTGCTACATGTCGCGCGGATGTTGTATACCTTACG 379
Qy 222 ValGluValLeuAspGlyArgValGlyValLeuGluAlaAlaGlyGluGluAlaArg 241
Db 378 GTGGAAGTCTCTCATGTTGGTGTGTACTGTCCTCAAGCGGCCCATTGAGGAGCGAGA 319
Qy 242 HisAspGlyArgLeuAlaHisLeuGlyArgAlaGluAspHisAlaLeuAlaValLeu 261
Db 318 CACGATGGAGCTCTTGCCACCCGAGAGGACCCAGCACCACCATGCGGTATGAGTTT 259
Qy 262 GlyArgAspThrGlnLeuArgValAlaAlaArgAlaHisLeuLeuAspHis 277
Db 258 GCGCGGACACTGAGCTGCGATTCCTCGCTGGGACAAAGTCTTCATCAT 211

RESULT 2
US-09-960-352-10273/c
; Sequence 10273, Application US/09960352
; Patent No. US2002013139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Bvatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112

```

```

; SEQ ID NO 10273
; LENGTH: 405
; TYPE: DNA
; ORGANISM: Bos taurus
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (3),(23),(31)
; OTHER INFORMATION: unsure at all n locations
; OTHER INFORMATION: Clone ID: 44-LIB3058-050-Q1-K1-C8
US-09-960-352-10273

Alignment Scores:
Pred. No.: 7,61e-29 Length: 405
Score: 397.00 Matches: 82
Percent Similarity: 81.98% Conservativeness: 9
Best Local Similarity: 73.87% Mismatches: 20
Query Match: 28.14% Indels: 0
DB: 10 Gaps: 0

US-09-709-103-1F4 (1-282) x US-09-960-352-10273 (1-405)
Qy 172 GluGlyValAlaValValGlnThrGluHisGlnAspGluAsnValSerCysGluAspGly 191
Db 404 GAAGGAGTCCGCGTGTACTGTCTGAACCTTAGGATGAACACGCTCTCGGTAAAGATGCA 345
Qy 192 GluAlaProHisGlyGlyGluArgValValAlaGlyArgValGluAspValGluLeuVal 211
Db 344 GAGGCGTCGATGCGCGGGAATGCGCGGTGCGGTACGTGTCGAGGATGTCAGCTGGTA 285
Qy 212 AspLeuAlaAlaAspGlyValGluLeuAlaValGluValLeuAspGlyArgArgValGly 231
Db 284 CCTCTCGCTCGAATCAGTAGAATCTGCGGTGGAAGTCTTCGATGTTGGCGTGTAGGC 225
Qy 232 ValLeuGluAlaAlaGlyGluGluAlaArgHisAspGlyArgLeuAlaHisLeuGlyArg 251
Db 224 GTCTCTGATGCGCGCGCTCAGGAAGCGGACACGATGGCGCTTGTCCACCTTGGACCA 165
Qy 252 AlaGluAspHisAlaLeuAlaValLeuGlyArgAspThrGlnLeuArgValAlaArg 271
Db 164 GCCGAGGACCATCGGTGTAGCAGTCTTGGCGGAGATCCTCACATCCGAGTCCGTCGG 105
Qy 272 AlaHisLeuLeuAspHisArgGlyGlnPheHis 282
Db 104 GCACATTTCTTCATCATCGCACAGTTTCAT 72

RESULT 3
US-09-778-963A-3/c
; Sequence 3, Application US/09778963A
; Patent No. US20020115172A1
; GENERAL INFORMATION:
; APPLICANT: NEELAM, Beena et al
; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL001112
; CURRENT APPLICATION NUMBER: US/09/778,963A
; CURRENT FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 11221
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-778-963A-3

Alignment Scores:
Pred. No.: 8,7e-25 Length: 11221
Score: 371.50 Matches: 97
Percent Similarity: 60.32% Conservativeness: 17
Best Local Similarity: 51.32% Mismatches: 62
Query Match: 26.33% Indels: 13
DB: 10 Gaps: 5

```

US-09-709-103-1f4 (1-282) x US-09-778-963A-3 (1-11221)

QY 4 Aapaaalaleuvalleu---Glyeuaalalaglaleuvalaaspval 22
 Db 8217 GATGGTGCACCTTCCCTCAGGAGGCTGCTCCGAGGAGCCTTGCTATGTA 8158
 QY 23 HlsagvalalavalTYralaglyProalarglulGlyCyshisaspalalysglYval 42
 Db 8157 CTGAGGCTGCTGACGCTGGGGCGCGGAGGCGAGACCATGCTAGGCGCTC 8098
 QY 43 AlaargvalalalalalalalalalaglYProglInglleuvalProglInarg 62
 Db 8097 ---CATCTCCTTACCGCGGCGGCTGAGAGGCGCTGGG----- 8062
 QY 63 leuvalaglInhlsvaAlalavalLeuhsAraspLeuvalalglvalTPalshis 82
 Db 8061 -----GTGGAAGCGCTCACCGTACTGCACGAGATCTTGCATGCGAGGCGGCTCAT 8008
 QY 83 leuvalaglYleuvalYhlaglYglularglaglulhlsleuvalglmalalalaval 102
 Db 8007 CTGCTGGGACGCTTGGCCATCTGTAAGACAGCTAGAACATCTGCTCACGTTGGTGT 7948
 QY 103 leuvalYArxAspLeuvalYvalalalaleuvalYvalalhlsglInleu 122
 Db 7947 CTCTTGGCGGACACCTCGAAGTAGGCGCAGTCTC---GTCCGCGGACACCGAGCTTC 7891
 QY 123 Aspleuvalaleuvalhlleuvalalglvalalal---ValThleuvalalalasp 141
 Db 7890 GGCCTCGGTGGGCGACCTGCGGCGGACAGCTGCTGCTGCTTCTTGGCCAGAT 7831
 QY 142 AspgInglYhlsvaAlalaleuvalYPhavalLeuvalalargleuvalYvalglu 161
 Db 7830 GACCATGGGACGCTCGCGGCTGCTTGGCTTCTTCAAGCAGACCTTACCTCAG 7771
 QY 162 AspleuvalProglProleuvalhlleuvalglYvalalalaleuvalYhlsgl 181
 Db 7770 GATGCTCTTGAAGGCGCTTGAACCTCATCGAAGGACCTCCGATATCCAGGCTGAACAC 7711
 QY 182 GlaspLeuvalleuvalserCyshlasp 190
 Db 7710 CAGGATGAACATCCCTGACAGGAG 7684

RESULT 4

US-09-864-761-21643/c
 / Sequence 21643, Application US/09864761
 / Patent No. US20020048763A1
 / GENERAL INFORMATION:
 / APPLICANT: Penn. Sharron G.
 / APPLICANT: Rank, David R.
 / APPLICANT: Hanzel, David K.
 / APPLICANT: Chen, Wensheng
 / TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 / FILE REFERENCE: Aesomica-X-1
 / CURRENT APPLICATION NUMBER: US/09/864,761
 / PRIOR FILING DATE: 2001-05-23
 / PRIOR APPLICATION NUMBER: US 60/180,312
 / PRIOR FILING DATE: 2000-02-04
 / PRIOR APPLICATION NUMBER: US 60/207,456
 / PRIOR FILING DATE: 2000-05-26
 / PRIOR APPLICATION NUMBER: US 09/632,366
 / PRIOR FILING DATE: 2000-10-04
 / PRIOR APPLICATION NUMBER: GB 24263,6
 / PRIOR FILING DATE: 2000-09-27
 / PRIOR APPLICATION NUMBER: US 60/236,359
 / PRIOR FILING DATE: 2001-01-30
 / PRIOR APPLICATION NUMBER: PCT/US01/00666
 / PRIOR FILING DATE: 2001-01-30
 / PRIOR APPLICATION NUMBER: PCT/US01/00667
 / PRIOR FILING DATE: 2001-01-30
 / PRIOR APPLICATION NUMBER: PCT/US01/00664
 / PRIOR FILING DATE: 2001-01-30
 / PRIOR APPLICATION NUMBER: PCT/US01/00669

/// PRIOR FILING DATE: 2001-01-30
 /// PRIOR APPLICATION NUMBER: PCT/US01/00665
 /// PRIOR FILING DATE: 2001-01-30
 /// PRIOR APPLICATION NUMBER: PCT/US01/00668
 /// PRIOR FILING DATE: 2001-01-30
 /// PRIOR APPLICATION NUMBER: PCT/US01/00663
 /// PRIOR FILING DATE: 2001-01-30
 /// PRIOR APPLICATION NUMBER: PCT/US01/00662
 /// PRIOR FILING DATE: 2001-01-30
 /// PRIOR APPLICATION NUMBER: PCT/US01/00661
 /// PRIOR FILING DATE: 2001-01-30
 /// PRIOR APPLICATION NUMBER: PCT/US01/00670
 /// PRIOR FILING DATE: 2000-09-21
 /// PRIOR APPLICATION NUMBER: US 60/234,687
 /// PRIOR FILING DATE: 2000-06-30
 /// PRIOR APPLICATION NUMBER: US 09/608,408
 /// PRIOR FILING DATE: 2001-01-29
 /// PRIOR APPLICATION NUMBER: US 09/774,203
 /// SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
 /// SEQ ID NO 21643
 /// LENGTH: 368
 /// TYPE: DNA
 /// ORGANISM: Homo sapiens
 /// FEATURE:
 /// OTHER INFORMATION: MAP TO AL022334.1
 /// OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.8
 /// OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.2
 /// OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3
 /// OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.6
 /// OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3
 /// OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.9
 /// OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.9
 /// OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.7
 /// OTHER INFORMATION: NT HIT: G11418034, EVALUATE 0.00e+00
 /// OTHER INFORMATION: SWISSPROT HIT: P13856, EVALUATE 8.00e-12
 /// OTHER INFORMATION: EST_HUMAN HIT: BE389944.1, EVALUATE 0.00e+00
 /// US-09-864-761-21643

Alignment Scores:

Pred. No.: 1,69e-12 Length: 368
 Score: 219.00 Matches: 50
 Percent Similarity: 64.44% Conservative: 8
 Best Local Similarity: 55.56% Mismatches: 32
 Query Match: 15.52% Indels: 0
 DB: 10 Gaps: 0

US-09-709-103-1f4 (1-282) x US-09-864-761-21643 (1-368)

QY 188 CysglaspGlygluAlarProhlglYglYgluArgvalalaglYargvalgluasp 207
 Db 367 TGTGAGATGAGACGCTGCGCATGGCGGGAAGGGTGTCCAGAGTATCCAGAT 308
 QY 208 ValgluLeuvalaspLeuvalalalaspGlyvalGluLeuvalalgluvalleuaspGly 227
 Db 307 GTGAGCTGTATCATGTCCGCGGAGTGTGTATCTTACCGGTGGAAGTCTTCATGT 248
 QY 228 ArgYargValglYvalleuvalalalaglYgluvalalarghlaspGlyargLeuvala 247
 Db 247 GGGGTGTACTGTCTCAAGCGGCCATTGAGGAGCAGACAGATGGAGCTTGTCC 188
 QY 248 HlsleuvalYarglaspLeuvalaspHlsalalalalaleuvalleuvalYargAspThglleu 267
 Db 187 CACCGAGAGCAGCAGCAGCAGCATGCGGTATGAGTTTGGGGGGGACACATGAGCGT 128
 QY 268 ArgYargValalarglhlhlleuvalaspHls 277
 Db 127 GCAGTCCCGCTGGACAAAGTCTTCATCAT 98

RESULT 5

us-09-709-103-1f4.p2n.rnpb

Tue Dec 31 15:27:52 2002

APPLICANT: Tao, Nengbing
 APPLICANT: Byatt, John C.
 APPLICANT: Mathialagan, Nagappan
 TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
 TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
 FILE REFERENCE: 16511.006/37-21(10298)C
 CURRENT APPLICATION NUMBER: US/09/960,352
 CURRENT FILING DATE: 2001-09-24
 NUMBER OF SEQ ID NOS: 15112
 SEQ ID NO 5830
 LENGTH: 454
 TYPE: DNA
 ORGANISM: Bos taurus
 FEATURE:
 NAME/KEY: unsure
 LOCATION: (181),(244),(385)
 OTHER INFORMATION: unsure at all n locations
 OTHER INFORMATION: Clone ID: 25-LIB3058-057-Q1-K1-G1
 US-09-960-352-5830

Alignment Scores:
 Pred. No.: 5,298-07 Length: 454
 Score: 160.50 Matches: 46
 Percent Similarity: 46.21% Conservative: 15
 Best Local Similarity: 34.85% Mismatches: 64
 Query Match: 11.37% Indels: 7
 Gaps: 2

US-09-709-103-1f4 (1-282) x US-09-960-352-5830 (1-454)

QY 126 LeuValHisLeuAlaValGluValAlaValThrLeuValAlaAlaAspGlnGlyHis 145
 Db 445 CTCCTGAGCTTATTCATTCCTGCGGAGGCTCTCAAGAGCGGCAATTCCTCGCA 388
 QY 146 ValHisValLeuLeuGlyPheValLeuGluAlaAargLeuGlyValGluAspLeuPro 165
 Db 388 CTNTGCGGGTGTACCTTTGAGTTG-----ACGGATCAGCTCAAA 347
 QY 166 GluProLeuHisLeuLeuGluGlyValAlaValAlaValGlnThrGluHisGlnAspGluAsn 185
 Db 346 GAGCGGCTTCAGCTCTCCAGGGTTGCTCTCTGCTAGATAGATAAACCAGGATGAAGC 287
 QY 186 ValSerCysGluAspGlyGluAlaProHisGlyGlyGluArgValAlaAlaGlyVal 205
 Db 286 GTGACCTGGCAATGCAAGCGCTGCAGCGCCCGGTAGCGCGCCCGGTGGTGC 227
 QY 206 GluAspValGluLeuValAspLeuAlaAspGlyValGluLeuAlaValGluValLeu 225
 Db 226 GGTGATGTGACGCGCACCGCTTTGCTGCTGAGCTAGCGCTGCGGTAGGTATCTTC 167
 QY 226 AspGlyArgArgValGlyValLeuGluAlaAlaGlyGluAlaAargHisAspGlyArg 245
 Db 166 CATGCTGGCAGGTAGCTCTCAGGAAGTTGCGCGCGCACCCACCTTTCACAGAGCGCT 107
 QY 246 LeuAlaHisLeuGlyArgAlaGluAspAspHisAla 257
 Db 106 CTTGCG 71

RESULT 7

US-09-476-242-20/c
 ; Sequence 20, Application US/09476242
 ; Patent No. US20020146683A1
 ; GENERAL INFORMATION:
 ; APPLICANT: BARNETT, Susan
 ; APPLICANT: MARTIN, Eric
 ; APPLICANT: HARTOG, Karin
 ; TITLE OF INVENTION: MODIFIED HIV ENV POLYPEPTIDES
 ; FILE REFERENCE: 1605.002
 ; CURRENT APPLICATION NUMBER: US/09/476,242
 ; CURRENT FILING DATE: 1999-12-30
 ; NUMBER OF SEQ ID NOS: 26
 ; SOFTWARE: Patentin Ver. 2.0
 ; SEQ ID NO 20

US-09-960-352-4253/c
 ; Sequence 4253, Application US/09960352
 ; Patent No. US20020137139A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Warren, Wesley C.
 ; APPLICANT: Tao, Nengbing
 ; APPLICANT: Byatt, John C.
 ; APPLICANT: Mathialagan, Nagappan
 TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
 TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
 FILE REFERENCE: 16511.006/37-21(10298)C
 CURRENT APPLICATION NUMBER: US/09/960,352
 CURRENT FILING DATE: 2001-09-24
 NUMBER OF SEQ ID NOS: 15112
 SEQ ID NO 4253
 LENGTH: 452
 TYPE: DNA
 ORGANISM: Bos taurus
 OTHER INFORMATION: Clone ID: 18-LIB34-086-Q1-E1-E5
 US-09-960-352-4253

Alignment Scores:
 Pred. No.: 2,258-07 Length: 452
 Score: 164.50 Matches: 56
 Percent Similarity: 46.58% Conservative: 19
 Best Local Similarity: 34.78% Mismatches: 73
 Query Match: 11.66% Indels: 13
 Gaps: 5

US-09-709-103-1f4 (1-282) x US-09-960-352-4253 (1-452)

QY 97 LeuValGlnAlaAlaValLeuLeuGlyArgAspLeuGluValGlyAlaAlaAlaLeuGly 115
 Db 447 CTCCTGAGCTTATTCATTCCTGCGGAGGCTCTCAAGAGCGGCAATTCCTCGCA 388
 QY 116 ValValAlaHisGlnLeuLeuAspLeuAlaValHisValLeuGluGlyPheValLeuAlaVal 135
 Db 387 CGCATAGGCGAGCCTTCCTT-----CTCACTCACCTCCGACGCGCTCTC---ATC 340
 QY 136 ThrLeuValAlaAlaAspGlnGlyHisValHisValLeuLeuGlyPheValLeuGlu 155
 Db 339 GCACCTGTTGCCACCAAGGATGACGGGCACTTTGCGGGTGTGTTACCTTTGAGTTG--- 283
 QY 156 AlaArgLeuGlyValGluAspLeuProGluProLeuHisLeuLeuGluGlyValAla 175
 Db 282 -----ACGGATCAGCTCATAGCGGCTTCAGCTCTCCAGGCTTGTGCTT 238
 QY 176 ValValGlnThrGluHisGlnAspGluAsnValSerCysGluAspGlyGluAlaProHis 195
 Db 237 CCTGGTATAGATAAACCAAGGATGAAGCGGTGATGTCAGCGCACCCCGCTTGTGCT 178
 QY 196 GlyGlyGluArgValAlaAlaGlyArgValGluAspValGluLeuValAspLeuAla 215
 Db 177 GCCCGGTAGCG 118
 QY 216 AspGlyValGluLeuAlaValGluValLeuAspGlyArgArgValGlyValLeuGluAla 235
 Db 117 GCAGCTAGCGCTTGGCGGTAGGTATCTTCGATGTCAGCGCACCCCGCTTGTGCT 58
 QY 236 AlaGlyGluGluAlaAargHisAspGlyArgLeuAlaHisLeuGlyArgAlaGluAsp 255
 Db 57 GCCCGCGCACCCACCTTTGCGCGCGCGCTTCTTCCCGCGCGCGCGCGCGCGCGCGCGCG 4
 QY 256 His 256
 Db 3 CAC 1

RESULT 6

US-09-960-352-5830/c
 ; Sequence 5830, Application US/09960352
 ; Patent No. US20020137139A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Warren, Wesley C.

Alignment Scores:	
Pred. No.:	4.52e+06
Score:	159.50
Percent Similarity:	35.6%
Best Local Similarity:	30.72%
Query Match:	11.30%
DB:	10
Length:	2322
Matches:	94
Conservative:	15
Mismatches:	116
Indels:	81
Gaps:	10

```

QY 12 LeuAlaAlaGlyValAGlyLeuLeuAlaAspValHisGluValAlaValTyrAlaGlyPro 31
Db .....:::
1086 GTTGGCTGCTGCACGCCGAGTCTGGCCGGGATGGGGGGGGCTGATCATGGCTTGGCCGCCA 10
QY 32 AlaATgGluGlyCysHisAspAlaLysGly-----ValAlaArgValAla 46
Db .....:::
1026 GCGGTGTATGTCTGCTGGATGGCGGAGGGGAGGTGATGGTCCCTGCTGTATTGG 96
QY 47 AlaAlaAlaAlaAlaAlaGlyProGlnGlnLeuLeu-----ValProGlnArgLeuLeu 64
Db .....:::
966 GCCATGTGCTGTGTCCACGGCTCTGTGAACACCTGGCTGTGGTCAGTAAAGAACTC 90
QY 65 ValGlnHisValAlaValLeuHisArgAspLeuAlaValGlnValTyrAlaHisLeuAla 84
Db .....:::
906 GCCGCCGACAGTTGAAGCTGTGCATCAACGATCTGGG----- 87

```

[illegible]

RESULT 8
US-09-476-242-18/c
; Sequence 18, Application US/09476242
; Patent No. US20020146683a1

```

1 / GENERAL INFORMATION:
2 /
3 / APPLICANT: BARRETT, Susan
4 / APPLICANT: HARBOG, Karin
5 / APPLICANT: MARTIN, Eric
6 / TITLE OF INVENTION: MODIFIED HIV ENV POLYPEPTIDES
7 / FILE REFERENCE: 1605.002
8 / CURRENT APPLICATION NUMBER: US/09/476,242
9 / CURRENT FILING DATE: 1999-12-30
10 / NUMBER OF SEQ ID NOS: 26
11 / SOFTWARE: PatentIn Ver. 2.0
12 / SEQ ID NO 18
13 / LENGTH: 2322
14 / TYPE: DNA
15 / ORGANISM: Artificial Sequence
16 / FEATURE:
17 /
18 / OTHER INFORMATION: Description of Artificial Sequence: Lem122-Seq199
19 / OTHER INFORMATION: Arg426-Gly431
20 / 09-476-242-18

```

Query NO.:	1	176-05
Score:	155	00
Percent Similarity:	33.00%	
Best Local Similarity:	30.25%	
Query Match:	10.99%	
DB:	10	
	Gaps:	16
	Matches:	2322
	Conservative:	103
	Mismatches:	16
	Indels:	129
		92
		16

US-09-709-103-1F4 (1-282) X US-09-476-242-18 (1-2322

QY	10	Length	LeuAlaAlaGlyAlaGlyLeuLeuAlaAspValHisGluValAla-----	26
Db	1251	CTTGCCCTTGCTGGGGCCACGCCACGGGCGCTGCATCTTCCACCACTTGTA	CTTGACAG	
QY	27	-----Val	TyrAlaGlyProAlaArgGluGlyCysHisAspAlaValGlyVal	42
Db	1191	CTCGGTGGGCCAGTTGTCGGGACATGCGCGCGCGGGCGGAGAAATCTCG	TGSGTGTGTT	113

43	Ala-----	-ArgValAlaAlaAlaAla	49
44			
Db	1131 GCTGATCTCTTGCCTGCCTGCGGGTCAAGCAGAGCCGGTGATGTCTCTCGAGCG	1072	

50 -----AlaAlaAlaGlyProGlnGlnLeu 57
1071 GATCTGCCCGGATGGGGGGGCGTACATGCCTTTCGCCCGCGCGGTTTATCATCTG 1012

1011 CTTGATGCGGCGAGGCGATGATGCGCTTGCTGTGTTGGGCCGATGCGTGTGTT 952

[illegible]

831 GCCGACTGGGCTGCAGCTTGT-----CAGCATCTGCTTCAGGGNGTGTTCCA 781

27 -----ValtyrAlaGlyProAlaArgGluGlyCysHisAspAlaLysGlyVal 42
1197 CTCGTCGGCAGTGTTCGGCATGTCGCCGCCGCCGGCGAAGATCTCGTGGTGT 1138
43 AlaArgValAlaAlaAlaAlaAlaAlaGlyProGlnGlnLeuValProGlnArg 62
1137 GCTGATCTC-----CTTGGCGCGTTCGGGGTTCAGCAGCAG----- 1102
63 LeuLeuValGlnHisValAlaValLeuHisArgAspLeuAla-----Val 77
1101 -----GCCGGTGATGTGCTGCTGCA-----CGGATCTGGCGCGGATGGGGGGCGTA 1051
78 GlnValTrpAlaHisLeu-----AlaGlyGlnLeu 87
1050 CATGGCCITGGCCACCTCTGCCAGCGGTTGATGATCTGCTTGATGCGCGCAGGGCAG--- 994
88 GlyHisGly----- 90
993 GGTGATGTTGCCCTGTTGTTGGGGCCGATGGGTTGTTCCAGGTGCTGTTGAACAG 934
91 -----GluGluArgAlaGluHisLeu----- 97
933 CTGGTCTCTTGCAGTAGAAGAACTCGCCGCCGCGAGTTGAAGCTGTGCATCAGATCTC 874
98 -----ValGlnAlaValLeuLeuGlyArgAsp-----LeuGluValGlyAla 113
873 GGGGTCCGCCCGCTCTCTCTGTTGAACACAGCATGCTCTTGTCCGAACTGGCGCTCGAG 814
114 LeuGlyValValAlaHisGlnLeuLeuAspLeuAlaValHisLeuAlaValGlnVal 133
813 CTTGGT-----CACGATCTGCTCAGGGTGTGTTCTTCCACTTCTCGCCGCTGATGTT 763
134 AlaValThrLeu-----ValAlaAlaAspAspGlnGlyHisValHis----- 147
762 CAGGTGGCGCTTGGCGGATGTTCGCCGATGATCTCGCCGTTGGCGGTAGAAAGCGCGCGCG 703
148 -----ValLeuLeuGlyPheValLeuGluAla----- 156
702 GCCGATGGTGATGCTCTTGGCGGTGTGTTGTTGGGCGGGTGCGATGATCTCCAGCT 643
157 -----ArgLeuGly-----ValGluAspLeuLeuProGluProLeu 168
642 CTCCTTCAGCTGCAGATGATGCTCTTGGCGTTGTCGGTGAAGTTCTCGTCGGGATCAC 583
169 HisLeuLeuGluGlyValAlaValValGlnThrGluHisGlnAspGluAsnValSerCys 188
582 CAGCCCTCTCTCGCCAGCGTTCGCTTCAG----- 553
189 GluAspGlyGluAlaProHisGlyGlyGluArgValAlaAlaGlyArgValGluAspVal 208
552 CAGCAGCTGGTGCTTCACACGGGGCGGATGCCGTG-----GGTGCATCGACGGTGCT 499
209 GluLeuValAspLeuAlaAspGlyValGluLeuAlaValGluValLeuAspGlyArg 228
498 CAC---GTTGGTGCAGGGCGCGCTCCGCTTGAATCTTGTGCTGCATCTCAGGATGGC 442
229 ArgValGlyValLeuGluAlaAlaGlyGluGluAlaArgHisAspGlyArgLeuAlaHis 248
441 GAACCCGCG-----GGGGCGCAGTAGTGGATGGGATGGGCTCGAAGCTCAC 394
249 LeuGlyArgAlaGluAspAspHisAlaLeuAlaValLeuGlyArgAspThrGlnLeuArg 268
393 CTTGGGCGAGCGCTGGTGTGATCACGCTGTTGTCACAGCTTTCACGCGGGCTTCAGGCTCTG 334
269 ValAlaArgAla-----HisLeuLeuAspHis 277
-----GCTGGTGGTATCTGCTCCACCAT 292

RESULT 10
US-09-372-529-1/c
; Sequence 1, Application US/09972529
; Patent No. US20020150916A1
; GENERAL INFORMATION:

[illegible]

```

RESULT 9
US-09-476-242-7/C
; Sequence 7, Application US/09476242
; Patent No. US20020146683A1
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: HARTOG, Karin
; APPLICANT: MARTIN, Eric
; TITLE OF INVENTION: MODIFIED HIV ENV POLYPEPTIDES
; FILE REFERENCE: 1605.002
; CURRENT APPLICATION NUMBER: US/09/476,242
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 2334
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Leu122-Ser199
; US-09-476-242-7

```

Alignment Scores:	1.18e-05	Length:	2334
Pred. No.:	155.00	Matches:	107
Score:	38.4%	Conservative:	15
Percent Similarity:	30.2%	Mismatches:	114
Best Local Similarity:	10.93%	Indels:	118
Query Match:	10	Gaps:	20
DB:			

US-09-709-103-1F4 (1-282) x US-09-476-242-7 (1-2334)

QY 10 LeuGlyLeuAlaAlaGlyAlaGlyLeuLeuAlaAspValHisGluValAla----- 26

DB 1257 CTTGGCCVTGTGGGGCCACGCCAGGGGCTCGATCTTCCACCTTGTTACTTGTTACAG 1198

APPLICANT: Millennium Pharmaceuticals, Inc.
 APPLICANT: Meyers, Rachel
 TITLE OF INVENTION: 47316, A NOVEL HUMAN G-PROTEIN AND USES
 FILE REFERENCE: 38155-20041.00
 CURRENT APPLICATION NUMBER: US/09/972, 529
 PRIOR FILING DATE: 2002-04-08
 PRIOR FILING DATE: 2000-10-05
 NUMBER OF SEQ ID NOS: 9
 SOFTWARE: FASTSEQ for Windows Version 4.0
 SEQ ID NO 1
 LENGTH: 1224
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (313)...(924)
 US-09-972-529-1

Alignment Scores:
 Pred. No.: 6.8e-06 Length: 1224
 Score: 154.00 Matches: 84
 Percent Similarity: 40.14% Conservative: 34
 Best Local Similarity: 28.57% Mismatches: 100
 Query Match: 10.91% Indels: 77
 DB: 10 Gaps: 14

US-09-709-103-1f4 (1-282) x US-09-972-529-1 (1-1224)

```

QY 2 AlaAspAvalAlaLeuValLeuGlyLeuAlaAlaGlyLeuAlaAlaAsp 20
Db 1068 GCTCCCTCTCCGCTGCTGCTTCTCAGCGGAGTGCTGCTCAGGCGCTCAGCGCGG 1029
QY 21 AspValHisGlu-----ValAlaValTyrAlaGlyProAla 32
Db 1028 GACCGTCACAGTTCATTTCCCGCGCTCCCGCCAGCCCGCCAGTACGAGCGCCGCC 969
QY 33 ArgGluGlyCysHisAspAlaValGlyValAlaArgValAlaAlaAlaAlaAla 52
Db 968 -----CTCCGCTCCGCGCTGCTCCGCTGCGAGCCGA 939
QY 53 GlyProGlnGlnLeuValProGlnArgLeuValGlnHisValAlaValLeuHis 72
Db 938 GGGCGCGCGAGG-----CTCACATGATGCGCGAGCGG-GTTGCGCGCG--CAG 895
QY 73 ArgAspLeuAlaValGlnValTyrAlaHisLeu-----AlaGlyGlnLeuGlyHisGly 90
Db 894 CGCGCCCTGGAGCGGAGCGGAGCGGCTGTCAGAGCGGCGGCGGCGAGCGCTTT 835
QY 91 GluGlnArgAlaGlnHisLeuValGlnAlaAlaValLeuGlyArgAspLeuGlnVal 110
Db 834 GAGCAGCTCGCTTAAGAGCAGCAGATGTGCGCATTTGCTGCGCGAGCATTCACAGTA 775
QY 111 GlyAlaAlaLeuGlyValValAlaHisGlnLeuLeuAspLeuAlaLeuValHisLeuAla 130
Db 774 ---GCGGCACTTCAGAGTCTTGGGTACAGGTGCGAGCATTCGCGGAGTTCAGCG 718
QY 131 ValGluValAlaValThrLeuValAlaAlaAspArgGlnHisVal-----146
Db 717 TCCGCGCTGGAGTCCCGCTTGTGCGCAGATGATGAGGCGCTGTCAGAGTTCCGAT 658
QY 147 HisValLeuLeuGlyPheValLeuGlnAlaAspLeuValGlnAspLeuProGln 166
Db 657 CACCCCTCGCTC-----CAGGATCTCTGCGCGAT 628
QY 167 ProLeuHisLeuLeuGlnGlyValAlaValAlaGlnThrGlnHisGlnAspGluAsnVal 186
Db 627 GGTCTTGAAGTACTAAG-----GCTGCAAGAGCAGAGATGTGTTAGACACAGAGTGA 574
QY 187 SerCysGluAspGlyGluAlaProHisGlyGly-----197
Db 573 GCGGTGGAACACTCCGAGTCCCTCTCAGAGAGGTGTGCGCCAGCTCTGAGGCGTATTGAC 514
  
```

```

QY 198 ---GluArgValAlaAlaGlyArgValGluAspValGlnLeuValAspLeuAlaAlaAsp 216
Db 513 AGGGAAGCGCGTGAAGGCTGGAAGATTCAGAGTCTGAGGCTGCGAGCGCTTCAT 454
QY 217 GlyValGluLeuAlaValGluValLeuAspGlyArgArgValGlyValLeuGlnAlaAla 236
Db 453 GAC-----GACAGCAGCAGGTA-----AAGCGCGCG 427
QY 237 Gly-----GluGlnAlaArgHisAspGly 244
Db 426 GCGCGCTGGTGGGAGCAGCAGACCTCGTGAACCTGTTGTAACAAGTGGCGCAGATGCG 367
QY 245 ArgLeuAlaHisLeuGlyArgAlaGluAspAspHisAlaIle 258
Db 366 ACTCTTGCCCACTCGCGCGCGCCAGCAGCGGACCGCGGTA 325
  
```

RESULT 11

US-09-788-654A-1/c
 Sequence 1, Application US/09788654A
 Patent No. US20020119920A1
 GENERAL INFORMATION:
 APPLICANT: YE, Jane et al.
 TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,
 TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE
 FILE REFERENCE: CL001143
 CURRENT APPLICATION NUMBER: US/09/788, 654A
 NUMBER OF SEQ ID NOS: 20
 SOFTWARE: FASTSEQ for Windows Version 4.0
 SEQ ID NO 1
 LENGTH: 3192
 TYPE: DNA
 ORGANISM: Human
 US-09-788-654A-1

Alignment Scores:
 Pred. No.: 2.12e-05 Length: 3192
 Score: 154.00 Matches: 84
 Percent Similarity: 40.14% Conservative: 34
 Best Local Similarity: 28.57% Mismatches: 100
 Query Match: 10.91% Indels: 77
 DB: 10 Gaps: 14

US-09-709-103-1f4 (1-282) x US-09-788-654A-1 (1-3192)

```

QY 2 AlaAspAvalAlaLeuValLeuGlyLeuAlaAlaGlyLeuAlaAlaAsp 20
Db 1103 GCTCCCTCTCCGCTGCTGCTTCTCAGCGGAGTGCTGCTCAGGCGCTCAGCGCGG 1044
QY 21 AspValHisGlu-----ValAlaValTyrAlaGlyProAla 32
Db 1043 GACCGTCACAGTTCATTTCCCGCGCTCCCGCCAGCCCGCCAGTACGAGCGCGCG 984
QY 33 ArgGluGlyCysHisAspAlaValGlyValAlaArgValAlaAlaAlaAlaAlaAla 52
Db 983 -----CTCCGCTCCGCGCTGCTCCGCTGCGAGCGGCGGCGGCGGCGGCGG 954
QY 53 GlyProGlnGlnLeuValProGlnArgLeuValGlnHisValAlaValLeuHis 72
Db 953 GGGCGCGCGAGG-----CTCACATGATGCGCGAGCGG-GTTGCGCGCG--CAG 910
QY 73 ArgAspLeuAlaValGlnValTyrAlaHisLeu-----AlaGlyGlnLeuGlyHisGly 90
Db 909 CGCGCCCTGGAGCGGAGCGGAGCGGCTGTCAGAGCGGCGGCGGCGGCGCTTT 850
QY 91 GluGlnArgAlaGlnHisLeuValGlnAlaAlaValLeuGlyArgAspLeuGlnVal 110
Db 849 GAGCAGCTCGCTTAAGAGCAGCAGATGTGCGCATTTGCTGCGCGAGCATTCACAGTA 790
QY 111 GlyAlaAlaLeuGlyValValAlaHisGlnLeuLeuAspLeuAlaLeuValHisLeuAla 130
  
```

Tue Dec 31 15:27:52 2002

```

Db 789 ---GCCGCACCTTCAGGCTTGTGCTACAGGTGCGACACGTTCCAGCGGGGATCACGCG 733
Qy 131 ValGluValAlaValThrLeuValAlaAlaAspAspGlnGlyHisVal----- 146
Db 732 TCCGCGTGCAGGTCCTCGCTTGTGTCACAGATGATGATGGGCGTCTCGAGGTTCCGAT 673
Qy 147 HisValLeuLeuGlyPheValLeuGluAlaArgLeuGlyValGluAspLeuLeuProGlu 166
Db 672 CACCCCTCGTCTC-----CAGGATCTGCTGGCGGAT 643
Qy 167 ProLeuHisLeuLeuGluGlyValAlaValAlaValGlnThrGluHisGlnAspGluAsnVal 186
Db 642 GGTCTTCAGCTACTCAA-----GCTGTCAAAGCAGCAGATGCTGTAGACACGAGTGTGA 589
Qy 187 SerCysGluAspGlyGluAlaProHisGlyGly----- 197
Db 588 GCGGTGGACACTCCGGAGTCCCTCGACGAGGTGTCTGCCCATCTCTCGAGCGGTATTGAC 529
Qy 198 ---GluArgValAlaValAlaGlyArgValGluAspValGluLeuValAlaAlaAsp 216
Db 528 AGGGAAGCGCTGATGGGTGGAAGTCGAGGATCTGGAGGTCTGGACGCTGGCGCTTCAT 469
Qy 217 GlyValGluLeuAlaValAlaValLeuAspGlyArgArgValGlyValLeuGluAlaAla 236
Db 468 GAC-----GACAGCAGGCGGTA-----AAGCGCGCG 442
Qy 237 Gly-----GluGluAlaArgHisAspGly 244
Db 441 GCGCGTGGTGGGACGCGACCTCGCTGAACCTGTTGTACAGAACTGGCGCAGATGGC 382
Qy 245 ArgLeuAlaHisLeuGlyArgAlaGluAspAspHisAlaIle 258
Db 381 ACTCTGCCACACCTCGCGCCCCCAGCAGCCGCCACCGGTA 340

RESULT 12
US-09-476-242-19/c
; Sequence 19, Application US/09476242
; Patent No. US20020146683A1
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: HARTOG, Karin
; APPLICANT: MARTIN, Eric
; TITLE OF INVENTION: MODIFIED HIV ENV POLYPEPTIDES
; FILE REFERENCE: 1605.002
; CURRENT APPLICATION NUMBER: US/09/476,242
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 19
; LENGTH: 2322
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Leu122-Ser199;
; OTHER INFORMATION: Arg426-Lys432
US-09-476-242-19

Alignment Scores:
Pred. No.: 1.79e-05 Length: 2322
Score: 153.00 Matches: 101
Percent Similarity: 35.78% Conservative: 21
Best Local Similarity: 29.62% Mismatches: 126
Query Match: 10.84% Indels: 94
DB: 15 Gaps: 15

US-09-709-103-1f4 (1-282) x US-09-476-242-19 (1-2322)

Qy 10 LeuGlyLeuAlaAlaGlyAlaGlyLeuLeuAlaAspValHisGluValAla----- 26
Db 1251 CTGGCGCTTGTGGGGCCAGCCAGCGGCTCGATCTTCCACCTTGTACTGTGTACAG 1192
Qy 27 -----ValTyrAlaGlyProAlaArgGluGlyCysHisAspAlaLysGlyVal 42

```

RESULT 13
 US-10-032-717-9/c
 ; Sequence 9, Application US/10032717
 ; Patent No. US20020151709A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Andre R. Abad
 ; APPLICANT: Nicholas B. Duck

```

APPLICANT: Xiang Feng
APPLICANT: Ronald D. Flannagan
APPLICANT: Theodore W. Kahn
APPLICANT: Lynn E. Sims
TITLE OF INVENTION: Genes Encoding No. US20020151709A1el Proteins with
FILE REFERENCE: 35716/237005
CURRENT APPLICATION NUMBER: US/10/032,717
PRIOR FILING DATE: 2001-10-23
PRIOR APPLICATION NUMBER: 60/242,838
NUMBER OF SEQ ID NOS: 48
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 9
LENGTH: 2010
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: CDS
LOCATION: (1)...(2010)
OTHER INFORMATION: Maize optimized Cry1218-1
LOCATION: (0)...(0)
OTHER INFORMATION: mol218-1
US-10-032-717-9

```

```

Alignment Scores:
Pred. No.: 6e-05 Length: 2010
Score: 146.50 Matches: 114
Percent Similarity: 36.53% Conservative: 23
Best Local Similarity: 30.40% Mismatches: 95
Query Match: 10.38% Indels: 143
Gaps: 25

```

US-09-709-103-1f4 (1-282) x US-10-032-717-9 (1-2010)

```

QY 7 LeuLeuValLeuGluLeuAlaAlaGluValLeuLeuAlaAspValHisGluValAla 26
DB 1155 CTGCTGAGAGTTGA---GCCGCGGAGACACGCGGTGTAGAGAT----- 1114
QY 27 ValTyzAlaGluProAlaArgGlu-----GlyCys 36
DB 1113 CTGCTGAGAGTTGA---GCCGCGGAGACACGCGGTGTAGAGAT----- 1114
QY 37 HisAspAlaLysGly-----ValAlaArgValAlaAlaAla----- 48
DB 1053 CACGCTGAGGCGCGGTATGATGCAACACGTCGCGCGCGGATGACGAGAGATCGAT 994
QY 49 ---AlaAlaAlaAlaGluProGlnGlnLeuLeuValProGlnArgLeuValGlnHis 67
DB 993 CAGCGCGAGAGCTGGGCGC-----CTTGTCTGATCCACAGA 961
QY 68 ValAlaValLeuHisArgAspLeuAlaValGlnValTyrAlaHisLeuAlaGluLeu 87
DB 960 GCCGATGAGAGACACGTTACCGCGCGGCGGCTGTATCCTCCGCGGTAGCTG 901
QY 88 Gly----- 88
DB 900 GGCCTTGCTCCATGGGAGGTGCGGCTGTCTAGTTGGGAGAGAGGCCACACGTC 841
QY 89 ---HisGlyGluGluArg-----AlaGlnHisLeuValGlnAlaAlaValLeuGlu 105
DB 840 GAGCAGCGGAGAGGTATCTCCGCGCGGAGCTGTGTAGTCCACCCACTGCTGGCGGA 781
QY 106 ArgAspLeuGluValGly-----AlaAlaLeuGlyValVal 117
DB 780 GGTGCTTGTAGAGCTTGGAGGCGGCTCATATACCACTTCAGGAGTGTGGAGTACTC 721
QY 118 AlaHisGlnLeuLeuAspLeuValHisLeuAlaValGluValAlaAla----- 135
DB 720 GCGCGGTAGAGCTT-----CATCTGGCGGCTGTAGTGTGTATGATGAT 679
QY 136 -----ThreValAlaAlaAsp----- 141

```

```

DB 678 GGTGTGACACAGCCCACTCTCCGCGAGATGAGAGGCTCTTGAGAGAGAGAGCTG 619
QY 142 -----AspGlnGlyHisVal-----HisValLeu 149
DB 618 GAGTTGGCGGCATGCGGTACACGGGTAGAGAGGACCTCGAAGTTGTATCACCGGGA 559
QY 150 LeuGlyPheValLeuGluAlaArgLeuGlyValGluAspLeuProGluProLeuHis 169
DB 558 GAGGCGCATGTACTG---GTTGAAGAGGAGTCTGAGAT---CTCGAAGCGGTTGCGCAC 505
QY 170 LeuLeuGluGly-----ValAlaValAlaGlnTrpGlnHisGluAspGluAsnVal 186
DB 504 GTGCGGAGAGGCGCGGAGCCCTTGAGGTTCTCTCCACTCTCCAGGCGGTAGGTA 445
QY 187 Ser---CysGluAspGlyGluAlaProHisGlyGlyGluArgValAlaGlyArgVal 205
DB 444 GAGCTGTGTATTGTGCGGAGGCTCTCCAGCTCGAGAGGCGCTTGTGCGGCGTACTC 385
QY 206 GluAspValGluLeuValAsp----- 212
DB 384 GCGCATCTT---CTGTTGATGAGCTCTCCACTGTCATGAAGATTTCACCTGGGA 328
QY 213 LeuAlaAlaAspGlyVal-----GluLeuAlaValGluValLeuAspGlyArg 228
DB 327 CTCTCGCGCGAGGCGCGAGAGATGTCATGATGAGCTGGGTGTAGAGGACACAGTGGGCC 268
QY 229 ArgValGlyValLeuGluAlaAlaGlyGluGluAlaArgHisAsp----- 243
DB 267 CACGAAGGACACCGCGAG---GCCGAGAGAGAGCTTGGCCAGATGTGATGGCGGCTT 211
QY 244 ---GlyArgLeuAla-----HisLeuGlyArgAlaGluAspHisAlaAla 259
DB 210 GCGGCGCTCTGCGCGGACAGACGACCTCGGAGAGC-----GGG 169
QY 260 ValLeuGlyArgAspTrpGlnLeuArgValAlaAlaArgAlaHisLeu 274
DB 168 GTACTCGGA-----GGCGTTCCGCGGACATCTT 139

```

RESULT 14

```

US-09-815-242-7899/c
Sequence 7899, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari L.
APPLICANT: Zyckind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
TITLE OF INVENTION: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
PRIOR FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0

```

SEQ ID NO 7899
LENGTH: 4074
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
FEATURE:
NAME/KEY: CDS
LOCATION: (1)...(4074)
US-09-815-242-7899

Alignment Scores:

Pred. No.: 0.000139 Length: 4074
Score: 146.50 Matches: 110
Percent Similarity: 33.63% Conservative: 43
Best Local Similarity: 24.18% Mismatches: 111
Query Match: 10.38% Indels: 191
DB: 10 Gaps: 24

US-09-709-103-1F4 (1-282) x US-09-815-242-7899 (1-4074)

QY 9 ValLeuGlyLeuAla---AlaGlyAlaGlyLeuLeuAlaAspValHisGluValAlaVal 27
Db 4017 GTTGAAGACTCGGGCATCGCGCTCCATCGGTGATCGCGTCCACGATGTTCTGTGA 3958
QY 28 TyrAlaGlyProAla-----ArgGluGlyCysHisAspAlaLysGlyVal 42
Db 3957 CATCTTGTTCGGCGCTTACGTCGTCGACTTACCGTTCAGCATTTCTCGAGGGTGA 3898
QY 43 AlaArgVal----- 45
Db 3897 CGCGCGCCATAGGCTTCAGCGGCCACACCTCATCTCACCAGAGCGCTGACCACCGAA 3838
QY 46 -----AlaAlaAlaAlaAlaAlaGlyPro 54
Db 3837 CTGTGCTTACCACCGCGGTGCTGGTAAACAGGCTGTACGAGCGGTGGAACGGGC 3778
QY 55 GlnGlnLeuLeuValProGlnArgLeuLeuValGlnHisValAlaValLeuHisArgAsp 74
Db 3777 GTGCATCTTGTGTCACACAGGTGTTTACGCTTGAGCATGTACATGTAGCCGACGGTGGT 3718
QY 75 LeuAlaValGlnValTrpAla-----HisLeuAla----- 84
Db 3717 CGAGCGCTCGAACTGTTGCCGTACGCGCTCGAAGACGATCTTGCGCGCTTTCGG 3658
QY 85 -----GlyGlnLeuGlyHisGly-----GluGluArg 93
Db 3657 CAGGTCGGCCAGCTTCAGCATGGCTTGATCTCGCGTTCCTTGGCGCATCGAACACGG 3598
QY 94 AlaGluHis----- 96
Db 3597 GGTGCCCATCGGTACGCCACCGCAGGTTCTTGGCCAGCGGAGGATCTCGTTGTCGCC 3538
QY 97 -----LeuValGln----- 99
Db 3537 CAGCTCGTCGAGTTTCTTCGCGACCGCGATCTCGTTGTAGATCTCTGTCAGGAACCTT 3478
QY 100 -----AlaAlaValLeu 103
Db 3477 AGCGAGTTCCGGACCTTTCGCGTGTCTTTCGAGCATGCGGTTGTATCTTTCGCCCGACGCC 3418
QY 104 LeuGlyArgAspLeuGluValGly----- 111
Db 3417 CTTGGCCCGGAGCCCGAGGTGGTTTCGAGGATCTGACCGATCTCATACCGACGCGTAC 3358
QY 112 -----AlaAlaLeuGlyValValAla----- 118
Db 3357 GCCCAGCGGTTTGAGGACGATGTGACCGCGGTGCGTTCGTCATCTCGCGCATGCTCTTC 3298
QY 119 -----HisGlnLeuLeuAspLeuAlaLeuValHisLeuAlaVal 131
Db 3297 CACCGGCATGATCACCGAGACCAACCCCTGTGTACCGTTCAGCGCGCGGCGCATCTTTCGCC 3238
QY 132 GluValAlaValThrLeu-----ValAlaAlaAspGlnGlyHisValHisVal 148
Db 132 GluValAlaValThrLeu-----ValAlaAlaAspGlnGlyHisValHisVal 148

Db 3237 CGGCTGGATCGACGCTTGATCGCCAGGTAGACCTTACGATCTTACGACCGCCGGAGC 3178
QY 149 LeuLeuGlyPheValLeuGluAlaArgLeuGlyValGluAspLeuLeuPro----- 165
Db 3177 CAGGTGCGTCCGCTCTGCTGACGCTTCTGTTCTCGAATCTTGTCTCCAGGAGCTG 3118
QY 165 ----- 165
Db 3117 GCGACGATCGCTGATGTAGGCTGGGCTTCTCCAGCTGTTCGTTACGGGCGTCTCGCC 3058
QY 166 -----GluProLeuHisLeuLeuGluGlyValAlaValValGlnThrGluHis 181
Db 3057 CATGCGCAGCTTGAACCACTCGCGCTCGAGACCGTTCGAGGTAGTCTCGTTCGATCTC 2998
QY 182 Gln-----AspGluAsnValSerCysGluAspGlyGluAlaProHisGlyGlyGlu 198
Db 2997 CGTGCCTTCTTTCAGCGCGCGCCACCTTTCAGCTTGGC-----ACCGACCGGCGGACG 2941
QY 199 ArgValValAlaGlyArgVal---GluAspValGluLeu-----ValAspLeuAlaAlaAsp 216
Db 2940 CAGCGCTCGAAGTTCGCGCTTCGACATCGGAACCTCTTCGTTAGGTCCTTTCGGAT 2881
QY 217 GlyValGluLeuAlaValGluValLeuAspGlyArgArgValGlyVal---LeuGluAla 235
Db 2880 CTGGTCGAGTTG-----CATCTTCTCGATGACAGCGCGCGGAATCGCGCTCGACGCC 2827
QY 236 ---AlaGlyGluGlu-----AlaArgHis---Asp 243
Db 2826 GTCGCGGTGGAAGACCTGAACTGATCGATGACGTACCTTGGTGGCGTCCGACACGCG 2767
QY 244 GlyArgLeuAlaHisLeuGlyArg---AlaGluAspAspHisAlaAlaValLeuGly 262
Db 2766 GGAGGTGCTCTCAGTCGGAGCGCTTCTCACCAGAGATCGCGCGCAGCAGCTTCTCTTC 2707
QY 263 ArgAspThrGlnLeuArgValAla---ArgAlaHisLeuLeuAsp 276
Db 2706 CGG---AGTCAGCTGGTCTCGCTTTCGGAGTGACCTTGGCGAC 2665

RESULT 15

US-09-972-529-3/c
Sequence 3, Application US/09972529
Patent No. US20020150916A1
GENERAL INFORMATION:
APPLICANT: Millennium Pharmaceuticals, Inc.
APPLICANT: Meyers, Rachel
TITLE OF INVENTION: 47316. A NOVEL HUMAN G-PROTEIN AND USES
TITLE OF INVENTION: THEREFOR
FILE REFERENCE: 38155-20041.00
CURRENT APPLICATION NUMBER: US/09/972,529
CURRENT FILING DATE: 2002-04-08
PRIOR APPLICATION NUMBER: US 60/237,716
PRIOR FILING DATE: 2000-10-05
NUMBER OF SEQ ID NOS: 9
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 612
TYPE: DNA
ORGANISM: Homo sapiens
US-09-972-529-3

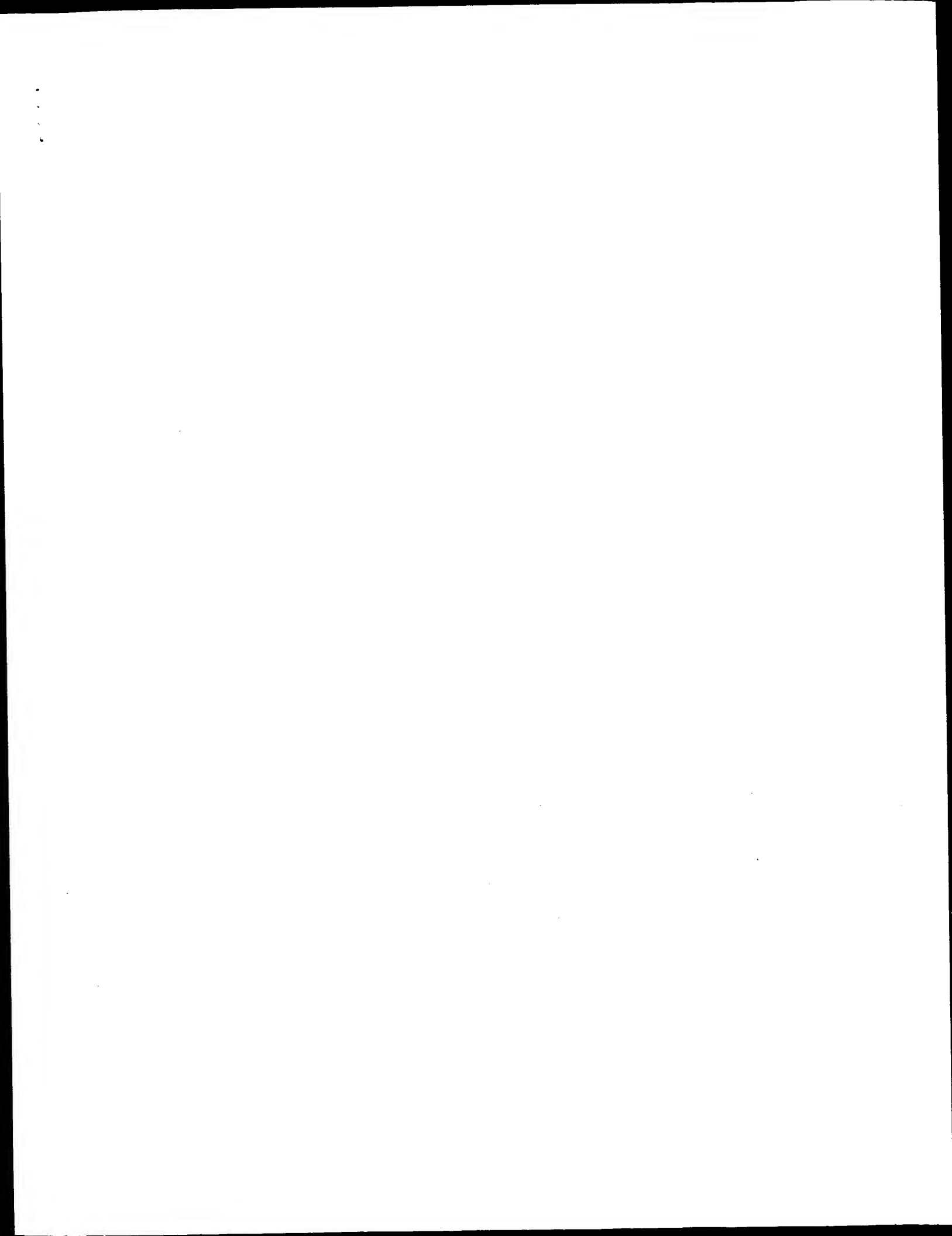
Alignment Scores:
Pred. No.: 3.08e-05 Length: 612
Score: 143.00 Matches: 66
Percent Similarity: 37.11% Conservative: 29
Best Local Similarity: 25.78% Mismatches: 75
Query Match: 10.13% Indels: 86
DB: 10 Gaps: 9

US-09-709-103-1F4 (1-282) x US-09-972-529-3 (1-612)

QY 29 AlaclyProAlaArgGluGlyCysHisAspAlaLysGlyValAlaArgValAlaAla 48
Db 600 GCAGCGTTCGGCGCGCGCCCTCGAAGCGCAGGCGAGCGTGCACGCTTGTGCAACG 541


```
QY 49 ALaAlaAlaAlaGlyProGlnGlnLeuValProGlnArgLeuLeuValGlnHisVal 68
Db 540 GCGGAGCGCGAGCTTGTGACGAGCTC----- 514
QY 69 AlaValLeuHisArgAspLeuAlaValGlnValTrpAlaHisLeuAlaGlyGlnLeuGly 88
Db 514 ----- 514
QY 89 HisGlyGlnGluArgAlaGlnHisLeuValGlnAlaAlaValLeuLeuGlyArgAspLeu 108
Db 513 -----GCTGAAGAGCAGCAGAGATGTGCCAGTTGCTTGTGCGAGCATTC 469
QY 109 GluValGlyAlaAlaLeuGlyValAlaHisGlnLeuLeuAspLeuAlaLeuValHis 128
Db 468 CACGTA--GCCGACCTTCAGGCTTTCGCTACCGAGGTGCACAGCTTCCAGCGCGGAT 412
QY 129 LeuAlaValGluValAlaValThrLeuValAlaAlaAspArgGlnGlyHisVal----- 146
Db 411 CACGGGTCCGGCGCTGCAGGTCCGCTTGTGCGCAGATGATGAGCGCTCTGAGGT 352
QY 147 -----HisValLeuLeuGlyPheValLeuGluAlaArgLeuGlyValGluAspLeu 164
Db 351 TCCGATCACCTCGCTC-----CAGGATCTGCTG 322
QY 165 ProGluProLeuHisLeuLeuGlnGlyValAlaValValGlnThrGlnHisGlnAspGlu 184
Db 321 GCGATGTGCTTGTGACGTACTCAA--GCTGTCAAGCAGCAGATGTCTGAGACAG 268
QY 185 AsnValSerCysGluAspGlyGluAlaProHisGlyGly----- 197
Db 267 GATGTAGGCGGTGACACTCCGAGTCCCTGCAGCAGGTGTCTGCCACTCTCGAGCGT 208
QY 198 -----GluArgValValAlaGlyArgValGluAspValGluLeuValAspLeuAla 214
Db 207 ATTCACAGGAGAGCGCTGATGGTGAAGATCGAGGATCTGAGGTCTGCACGTGCC 148
QY 215 AlaAspGlyValGluLeuAlaValGluValLeuAspGlyArgArgValGlyValLeuGlu 234
Db 147 GTTCATGAC-----GACAGCAGGCGAGTA-----AAG 121
QY 235 AlaAlaGly-----GluGlnAlaArgHis 242
Db 120 GCGGCGGCGCGGTGTGGGAGCGCAGACTCGCTGACTGTTGACAGAACTGCGGCAC 61
QY 243 AspGlyArgLeuAlaHisLeuGlyArgAlaGluAspAspHisAlaIle 258
Db 60 GATGGCACTTGTCCACACTCTGCGCCCCCAGCAGCGCCACCGGTA 13
```

Search completed: December 30, 2002, 18:07:35
Job time : 99 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: December 30, 2002, 16:16:46 ; Search time 77 Seconds

(without alignments)
1123.154 Million cell updates/sec

Title: US-09-709-103-1f4

Perfect score: 1411

Sequence: 1 LADDAALIVIGLAGAGLIA.....RDTQURVARAILDHRGQFH 282

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Zgapop 6.0 , Zgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 441362 segs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seg length: 0
Maximum DB seg length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame_plus.model -DEV=xlp
-O=/cgn2_1/USPTO/spool/STULIYANF4/runat 30122002.144047.21529/app.query.fasta_1.455
-DB=Issued Patents NA -OPMT=fastseq -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human4.0.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTPMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-NO XLPXY -NO MMAP -LARGEDBTRY -NEG SCORES=0 -WAIT -LONGLOG -DEV_TIMOUT=3
-MARK_TIMOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-GAPOP=10 -YGAPEXT=0.5 -DELDP=6 -DELEXT=7

Database:

1: Issued Patents NA.*
2: /cgn2_6/prodata/1/ina/5A_COMB.seg:*
3: /cgn2_6/prodata/1/ina/5B_COMB.seg:*
4: /cgn2_6/prodata/1/ina/5A_COMB.seg:*
5: /cgn2_6/prodata/1/ina/5B_COMB.seg:*
6: /cgn2_6/prodata/1/ina/5A_COMB.seg:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	1411	100.0	1841	4	US-09-053-374A-1
C 2	1348	95.5	3986	4	US-09-053-374A-1
C 3	963.5	68.3	1689	4	US-09-053-374A-4
C 4	911	64.6	3079	4	US-09-053-374A-6
C 5	166	11.8	3468	1	US-07-951-715A-2
C 6	166	11.8	3468	2	US-08-459-448A-2
C 7	166	11.8	3468	3	US-08-459-448A-2
C 8	166	11.8	3468	4	US-08-459-448A-2
C 9	166	11.8	3468	3	US-08-459-448A-2
C 10	166	11.8	3468	3	US-08-459-448A-2
C 11	166	11.8	3468	4	US-09-053-374A-3
C 12	159	11.3	3468	1	US-07-951-715A-4

C 13	159	11.3	3468	2	US-08-459-448A-4	Sequence 4, Appl1
C 14	159	11.3	3468	3	US-08-459-448A-4	Sequence 4, Appl1
C 15	159	11.3	3468	3	US-08-459-448A-4	Sequence 4, Appl1
C 16	159	11.3	3468	3	US-08-459-448A-4	Sequence 4, Appl1
C 17	159	11.3	3468	3	US-08-459-448A-4	Sequence 4, Appl1
C 18	159	11.3	3468	4	US-09-053-374A-5	Sequence 5, Appl1
C 19	153.5	10.9	1227	4	US-09-074-912-4	Sequence 4, Appl1
C 20	153.5	10.9	1227	4	US-09-074-912-3	Sequence 4, Appl1
C 21	153.5	10.9	1227	4	US-09-074-912-3	Sequence 4, Appl1
C 22	153.5	10.9	1227	4	US-09-074-912-3	Sequence 4, Appl1
C 23	153.5	10.9	1227	4	US-09-074-912-3	Sequence 4, Appl1
C 24	153.5	10.9	1227	4	US-09-074-912-3	Sequence 4, Appl1
C 25	153.5	10.9	1227	4	US-09-074-912-3	Sequence 4, Appl1
C 26	153.5	10.9	1227	4	US-09-074-912-3	Sequence 4, Appl1
C 27	153.5	10.9	1227	4	US-09-074-912-3	Sequence 4, Appl1
C 28	153.5	10.9	1227	4	US-09-074-912-3	Sequence 4, Appl1
C 29	153.5	10.9	1227	4	US-09-074-912-3	Sequence 4, Appl1
C 30	152	10.8	3546	1	US-07-951-715A-12	Sequence 12, Appl1
C 31	152	10.8	3546	2	US-08-459-448A-12	Sequence 12, Appl1
C 32	152	10.8	3546	3	US-08-459-448A-12	Sequence 12, Appl1
C 33	152	10.8	3546	3	US-08-459-448A-12	Sequence 12, Appl1
C 34	152	10.8	3546	3	US-08-459-448A-12	Sequence 12, Appl1
C 35	151	10.7	1848	4	US-09-547-422-12	Sequence 12, Appl1
C 36	151	10.7	1848	4	US-09-547-422-12	Sequence 12, Appl1
C 37	151	10.7	1848	4	US-09-547-422-12	Sequence 12, Appl1
C 38	150	10.6	3546	1	US-07-951-715A-14	Sequence 14, Appl1
C 39	150	10.6	3546	2	US-08-459-448A-14	Sequence 14, Appl1
C 40	150	10.6	3546	3	US-08-459-448A-14	Sequence 14, Appl1
C 41	150	10.6	3546	3	US-08-459-448A-14	Sequence 14, Appl1
C 42	150	10.6	3546	3	US-08-459-448A-14	Sequence 14, Appl1
C 43	144.5	10.2	1947	1	US-07-951-715A-3	Sequence 3, Appl1
C 44	144.5	10.2	1947	1	US-07-951-715A-3	Sequence 3, Appl1
C 45	144.5	10.2	1947	2	US-08-459-448A-3	Sequence 3, Appl1

ALIGNMENTS

RESULT 1
US-09-053-374A-1/c
; Sequence 1, Application US/09053374A
; Patent No. 6462177
GENERAL INFORMATION:
; APPLICANT: YEN, KWANG-MU
; TITLE OF INVENTION: MAMMALIAN BLOOD LOSS-INDUCED GENE, KD312
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESS: AMGEN INC.
; STREET: ONE AMGEN CENTER DRIVE
; CITY: THOUSAND OAKS
; STATE: CA
; COUNTRY: US
; ZIP: 91320
COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/053,374A
; FILING DATE:
; CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
; NAME: COOK, ROBERT R.
; REGISTRATION NUMBER: 31,602
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1841 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULAR TYPE: cDNA
; FEATURE:

us-09-709-103-1f4.p2n.rni

Tue Dec 31 15:27:52 2002

NAME/KEY: CDS
LOCATION: 255...1097
US-09-053-374A-1

Alignment Scores: 4,47e-126 Length: 1841
Pred. No.: 1411.00 Matches: 282
Score: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-709-103-1F4 (1-282) x US-09-053-374A-1 (1-1841)

QY 1 LeuAlaAspAlaAlaLeuValLeuGlyLeuAlaAlaGlyAlaGlyLeuLeuAla 20
DB 1100 CTAGCTGATGACGACGCTCTTGTCTTGGCTTGGCTGCGCGCTGCGCTTCTCGCG 1041
QY 21 AspValHisGluValAlaValTyrAlaGlyProAlaArgGluGlyCysHisAspAlaLys 40
DB 1040 GATGTACATGAGTCTGTGTACGCTGGCGCGCGCGAAGGGTGCACGATGCCAAA 981
QY 41 GlyValAlaArgValAlaAlaAlaAlaAlaAlaAlaGlyProGlnGlnLeuLeuValPro 60
DB 980 GCGCTGCGCGGCTGCGCGCGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCT 921
QY 61 GlnArgLeuLeuValGlnHisValAlaValLeuHisArgAspLeuAlaValGlnValTrp 80
DB 920 CAGGCTTCTTGTGCGACGCTGCGAGTACTGCGACCGAGACTTGGCGGTGAGTCTGG 861
QY 81 AlaHisLeuAlaGlyGlnLeuGlyHisGlyGluGluArgAlaGluHisLeuValGlnAla 100
DB 860 GCTCATCTCGCTGGCGAGCTTGGCCATGGCGAAGCGCGCGCGCGCGCGCGCGCGCT 801
QY 101 AlaValLeuLeuGlyArgAspLeuGluValGlyAlaAlaLeuGlyValValAlaHisGln 120
DB 800 GCTGTTCTTCTTGGCGAGACTTCTGAGTGGCGCGCGCGCGCGCGCGCGCGCGCGCT 741
QY 121 LeuLeuAspLeuAlaLeuValHisLeuAlaValGluValAlaValThrLeuValAlaAla 140
DB 740 CTGCTCGACTCTCGCGCTGCTCCACTCGCGGTAGAGTCTGCGCTGCGCTGCGCGCG 681
QY 141 AspAspGlnGlyHisValHisValLeuLeuGlyPheValLeuGluAlaLeuGlyVal 160
DB 680 GATGACCGAGGCGCGCTCCAGCTTCTCTTGGTCTTCTTCTGAGGCGAGACTTGGTCT 621
QY 161 GluAspLeuLeuProGluProLeuHisLeuLeuGlyValAlaValAlaValGlnThrGlu 180
DB 620 GAGGATCTGCTGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 561
QY 181 HisGlnAspGluAsnValSerCysGluAspGlyGluAlaProHisGlyGlyGluArgVal 200
DB 560 CACGAGTGAACAGTCTCTGCTGAGGATGAGAGCGCGCGCGCGCGCGCGCGCGCGGT 501
QY 201 ValAlaGlyArgValGluAspValGluLeuValAspLeuAlaAlaAspGlyValGluLeu 220
DB 500 GTTTCGCGAGCTGCTGAGAGTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCT 441
QY 221 AlaValGluValLeuAspGlyArgValGlyValLeuGluAlaAlaGlyGluAlaAla 240
DB 440 GCGGTGGAAGTCTCTGATGTAGGCTGTAGGCTGTAGGCTGTAGGCTGTAGGCTGTAG 381
QY 241 ArgHisAspGlyArgLeuAlaHisLeuGlyArgAlaGluAspHisAlaAlaVal 260
DB 380 CGACACGATGCGCTCTTGGCCACCTTGGACCGCGCGCGCGCGCGCGCGCGCGCGCT 321
QY 261 LeuGlyArgAspThrGlnLeuArgValAlaAlaArgAlaHisLeuLeuAspHisArgGlyGln 280
DB 320 CTTGCGCGGAGTACTCAGCTCCGAGTCCGCTCGGCGCATCTTCTTGTATCATCGCGCG 261
QY 281 PheHis 282
DB 260 TTTTCAT 255

RESULT 2

US-09-053-374A-3/c
Sequence 3, Application US/09053374A
Patent No. 6462177
GENERAL INFORMATION:
APPLICANT: YEN, KWANG-MU
TITLE OF INVENTION: MAMMALIAN BLOOD LOSS-INDUCED GENE, KD312
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: AMGEN INC.
STREET: ONE AMGEN CENTER DRIVE
CITY: THOUSAND OAKS
STATE: CA
COUNTRY: US
ZIP: 91320
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA: US/09/053,374A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: COOK, ROBERT R.
REGISTRATION NUMBER: 31,602
REFERENCE/DOCKET NUMBER: A-514
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3986 base pairs
TYPE: nucleic acid
STRADEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-09-053-374A-3

Alignment Scores: Length: 3986
Pred. No.: 138e-119 Matches: 282
Score: 1348.00 Conservatives: 0
Percent Similarity: 79.89% Mismatches: 0
Best Local Similarity: 79.89% Indels: 71
Query Match: 95.54% Gaps: 1
DB: 4

US-09-709-103-1F4 (1-282) x US-09-053-374A-3 (1-3986)

QY 1 LeuAlaAspAlaAlaLeuLeuValLeuGlyLeuAlaAlaGlyAlaGlyLeuLeuAla 20
DB 1830 CTAGCTGATGACGACGCTCTTGTCTTGGCTTGGCTGCGCGCTGCGCTTCTCGCG 1771
QY 21 AspValHisGluValAlaValTyrAlaGlyProAlaArgGluGlyCysHisAspAlaLys 40
DB 1770 GATGTACATGAGTCTGTGTGTAGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCT 1711
QY 41 GlyValAlaArgValAlaAlaAlaAlaAlaAlaAlaGlyProGlnGlnLeuLeuValPro 60
DB 1710 GCGGTGCGCGCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1651
QY 61 GlnArgLeuLeuValGlnHisValAlaValLeuHisArgAspLeuAlaValGlnValTrp 80
DB 1650 CAGCGCTCTTGTGCGACGCTGCGAGTACTGCGCGCGCGCGCGCGCGCGCGCGCT 1591
QY 81 AlaHisLeuAlaGlyGlnLeuGlyHisGlyGluGluArgAlaGluHisLeuValGlnAla 100
DB 1590 GCTCATCTGCTGGCGAGCTTGGCCATGCGAAGAGCGCGCGCGCGCGCGCGCT 1531
QY 101 AlaValLeuLeuGlyArgAspLeuGluValGlyAlaAlaLeuGlyValAlaHisGln 120
DB 1530 GCTGTTCTTCTTGGCGAGTCTCGAAGTGGCGCGCGCGCTGGGGGTCTGCGCCACCG 1471
QY 121 LeuLeuAspLeuAlaLeuValHisLeuAlaValGluValAlaValThrLeuValAlaAla 140

```

Db 1470 CTGCTCGATCTCGGCTGCTGCACCTCGCGTAGAAGTCCGCTCACTTGTGCGCA 1411
Qy 141 ASPASpGlnGlyHisValHisValLeuLeuGlyPheValLeuGlnAlaArgLeuGlyVal 160
Db 1410 GATGACCGAGGGGACGCTCCACCTTCTTCTTGTGTTGTTGAGGCAAGACTTGTGTC 1351
Qy 161 GluAspLeuLeuProGluProLeuHisLeuLeuGluGlyValAlaValAlaGlnThrGlu 180
Db 1350 GAGGATCTGCTGCTGAGCGCTGCACCTCTCGAAGAGTCCGCTTCTCCAGACTGAA 1291
Qy 181 HisGlnAspGlnAsnValSer 187
Db 1290 CACCAAGATGAAACCTCTCTTCTTCTGAGGGGGCAACAGAGCAAGAAAGAGTGAAGG 1231
Qy 187 187
Db 1230 TGGCCCGCCAGGGGACAAAGTCGGGCTGACTTGTAGCGCGCGGGGGCGGCTGAGCTTCTC 1171
Qy 187 187
Db 1170 TAAAGCCGAAAGCGCACTCTGCGCGGCGAGCAGCGGAGTCCGCTGCGGGGGCGCTG 1111
Qy 188 188
Db 1110 GCCGAGGCTTCCCGCCCTTCCCTCCGCACTGCGCGCGCGCGCGCTCACTGTGTGAGA 1051
Qy 190 PClyGlnAlaProHisGlyGlyGlnValAlaGlyValArgValGlnAspValGlnLeu 210
Db 1050 TGGAGAGGCGCGCGCGGCGGGAACGGGTGTGCGGAGCTGTCCAGAGATGTCAGACT 991
Qy 210 uValAspLeuAlaAlaAspGlyValGlnLeuAlaValGlnValLeuAspGlyValArgVal 230
Db 990 GGTAGACCTCCGCGCGGATGAGTGAAGTCTTGGGTGAAGTCTTCCATGTGTGAGGCTGT 931
Qy 230 GlnValLeuGlnAlaAlaGlyGlnGlnAlaArgHisAspGlyValArgLeuAlaHisLeuGln 250
Db 930 AGGGCTCTCCAAACGGCGCGGTGAGGAAGCGGACAGATGCGCTTCCCACTTGG 871
Qy 250 ValArgAlaGlnAspAspHisAlaAlaValLeuGlnValArgAspThrGlnLeuArgValAla 270
Db 870 ACGAGCGGAGGATGACATGCGATGAGCATGCTTGTGCGCGGATATCAGCTCCGAGTCCG 811
Qy 270 ValArgAlaHisLeuLeuAspHisAspGlyGlnPheHis 282
Db 810 TCGGCGACATCTTCTGATCATCGCGGCGAGTTTCAI 774

```

RESULT 3

```

US-09-053-374A-4/C
; Sequence 4, Application US/09053374A
; Patent No. 6462177
; GENERAL INFORMATION:
; APPLICANT: YEN, KWANG-MU
; TITLE OF INVENTION: MAMMALIAN BLOOD LOSS-INDUCED GENE, KD312
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSER: AMGEN INC.
; STREET: ONE AMGEN CENTER DRIVE
; CITY: THOUSAND OAKS
; STATE: CA
; COUNTRY: US
; ZIP: 91320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/053.374A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: COOK, ROBERT R.

```

```

;
; REGISTRATION NUMBER: 31,602
; REFERENCE/DOCKET NUMBER: A-514
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1689 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 132..971
; US-09-053-374A-4

Alignment Scores:
Pred. No.: 2,81e-83
Score: 963.50
Percent Similarity: 79.79%
Best Local Similarity: 70.92%
Query Match: 68.28%
DB: 4 Gaps: 1

US-09-709-103-1f4 (1-282) x US-09-053-374A-4 (1-1689)

Qy 1 LeuAlaAspAspAlaAlaLeuLeuValLeuAlaAlaGlyLeuAla 20
Db 974 CTACTGATGACAGAGCGCTCTTGTCTTACGCTGCTGACACTGTTTTCAGG 915
Qy 21 AspValHisGlnValAlaValAlaValAlaGlyProAlaArgGlnGlyValHisAspAlaVal 40
Db 914 AATGATCATGAGGTCGCTATGACAGCTAGCTTGCAGCAAGGGGCCAAGATGCCAA 855
Qy 41 GlyValAlaArgValAlaAlaAlaAlaAlaAlaAlaGlyProGlnGlnLeuValPro 60
Db 854 GGCATCTCC--GTGGTGGCGCCCGCCCGCTGCGCGCGAGAGAGCTTTCTTGTCT 798
Qy 61 GlnArgLeuLeuValGlnHisValAlaValLeuHisArgAspLeuAlaValGlnValTrp 80
Db 797 CAGAGCTTTTGTGACGACGTCACAGTACTGACAGACACTTGGCGTGCAGTCAAG 738
Qy 81 AlaHisLeuAlaGlyGlnLeuGlnHisGlyGlnGlnAlaGlnValAlaHisGln 100
Db 737 GCTCATCTCCCTGAGGAGCTTGGCGCAAGGCGAGGAGGAGGAGGAGGAGGAGGAGG 678
Qy 101 AlaValLeuLeuGlyValArgAspLeuGlnValAlaAlaLeuGlyValAlaHisGln 120
Db 677 GCTATCTTCTTGGCGGAGATCTCGAAGTGAAGCACAGCTGAGGTCATCGCCACAG 618
Qy 121 LeuLeuAspLeuAlaLeuValHisLeuAlaValAlaValAlaValThrLeuValAlaAla 140
Db 617 CTGCTCAATCTCCGCTGCTCCACTTCCGGTGAAGTCCCGGTCCCTTGTTCACGCA 558
Qy 141 ASPASpGlnGlyHisValHisValLeuLeuGlyPheValLeuGlnAlaArgLeuGlyVal 160
Db 557 AATGACCAAGCGGACGTCACATCTCTTGTGTTGTTGTTGAGACAGACTTGGTTC 498
Qy 161 GluAspLeuLeuProGluProLeuHisLeuLeuGlnGlyValAlaValAlaGlnThrGlu 180
Db 497 TAGGATCTGCTGTTGAGCTTTCGACCTTCCAGAGAGTCCGCTTGTCTAAGCTGAA 438
Qy 181 HisGlnAspGlnAsnValSerCysGlnAspGlyGlnAlaProHisGlyGlnGlnVal 200
Db 437 CACCAAGATGAAACCTCTCTGAGGATGAGAGGCGCGGAGTGGCGGAAACGAGT 378
Qy 201 ValAlaGlyValArgValAlaGlnAspValGlnLeuValAspLeuAlaHisAspGlyValGlnLeu 220
Db 377 ATTGCAGATGTGTCCAGTATGTCACCTGTTGACTTCCCGCGGATCCAGTAAACTT 318
Qy 221 AlaValGlnValAlaLeuAspGlyValArgValGlyValLeuGlnAlaAlaGlyGlnGlnAla 240
Db 317 TCGGTGGAAGCTTCAATGATGAGGGGTGAAGGCTCTCGAAGCGGCCGTGAGAGAGCG 258
Qy 241 ArgHisAspGlyValArgLeuAlaHisLeuGlnValArgAlaGlnAspAspHisAlaValAlaVal 260

```


Alignment Scores:	
Pred. No.:	
Score:	1,72e-06
Percent Similarity:	166.00
Best Local Similarity:	38.43%
Query Match:	29.97%
DB:	11.76%
	1
	Gaps:
	16

QY	1	LeuAlaAspAspAlaAlaLeuLeuValLeuGlyL-	-----LeuAla	13
Db	2304	GTAGCGGATCCAGTAGATCTCCAGGCTCCGCTCTGCTGATGATGACGGCGACGTGGTA		2244
QY	14	AlaGlyValAlaGlyLeuLeuAlaAspValHisGluValAlaValIYTrAlaGlyProAlaIatg		33
Db	2244	GCGGGTGTAGGCTTCACGCTTCTGCTGCGATCTTCGTGACAGGTAGTGGAGGTACGA		2185
QY	34	-----GluGlyCysHisAspAlaValuGlyValI		42
Db	2184	CTCGTGAAGGTCGCCAGACGGTCAAGTAGTCTCTTGAACAGTCGTGCGCGCCCTG		2125
QY	43	---AlaArgValAlaAlaAlaAlaAlaAlaAlaAlaGlyProGlnGlnLeuLeuValProGln	61	
Db	2124	GATGGTAGTGTGGGTGCTGGCCGCGCAGCGCGGCTCCAGCTGGGGGTATATGCCCGCGGAA		2065

QY 62 ArgLeuLeuValGlnHisValAlaValLeuNH₂SAARGSPRLLeuAlaValGlnValITPrAla 81
 Db 2064 GTTGGGCTCTCCAGCAGAGTTGGCGCTGCTGCTCAGAGCGCTTGGCGCTC-----CTT 201
 QY 82 HisLeuAlaGlyGlnLeuGlyHisIsgIyGlnGlnAArgIuHisIleLeuValGlnAlaAla 101
 Db 2013 CACCTTCTCGCTCAGCTC-----CTTCTTCTCGTCCAGGACAGA 1975
 QY 102 ValLeuLeuGlyAArgAspLeuGlnValGlnAlaAlaLeuGlnValValAlaHisGlnLeu 121
 Db 1974 CTC-----GTCCCTCAGGACACTCCACAGGTGTCTCAC---CTG 1939
 QY 122 LeuAspLeuAlaLeuValHisIleuAlaValGlnValAlaValHisIleuValAlaAlaAsp 141
 Db 1938 GTCCATGTGGTAGTCCGTCACTGCGCTTACG-----GCCATTTGGTTCCTCTGCT 1885
 QY 142 AspGln-----GlnHisValHisValLeuLeuGlyPheValLeuGln 155
 Db 1884 GAACAGCTGCTTACAGCGCTTCTGGGCGCGCTCCAGGTGACTCGGC-----CTGAA 1831
 QY 156 AlaArgLeuGlyValGlnAspLeuLeuProGlnProLeuHisIleuLeuGlnGlyValAla 175
 Db 1830 GGTCACTCTCGGCGGACAGAACTGATGCGGTGATGTACACTGTTGGCCGCTTGAA 1771
 QY 176 ValVal-----GlnThrGlnHisGlnAspGlnAsnValSerCysGlnAspGlyValAla 193
 Db 1770 CAGGTGGGCGCTCAAGGTGAACACGCTGCTGCGCTTGCTGAAGTTGAAGGCGGTGGAA 1711
 QY 194 ProHisGlyGlyGlnAArgVal-----ValAlaGlyAArgValGln-----Asp 207
 Db 1710 GCCCAAGCGCGAAGCTGCCGCTGTCAGAGTTGCTGCCCTGCTCATGTTGCGCTTGA 1651
 QY 208 ValGlnLeuValAspLeuAlaAla-----AspGlyValGlnLeuAlaVal----- 222
 Db 1650 GTTGGCCGTGTTATGGGCGCGCGCTGCATGCTGTGTGGAACTGCAGGTGGTGTGCT 1591
 QY 223 GlnValIleAspGlyAArgAArgValGlnValLeuGlnAlaAlaGlyGlnGlnAlaAlaGHis 242
 Db 1590 GCGCGAAGCGATGGCACGCGGTAGCGCTGGCTCAGGGGGCGGTATGTTACCGCGAG 1531
 QY 243 AspGlyAArgLeuAlaHisIleuGlyAArgAlaGlnAspHisAlaIleAlaValLeuGly 262
 Db 1530 GGTGCTGATGTCGCGCGGCTGTGTGCGCGAGAT----- 1495
 QY 263 ArgAspThrGlnLeuAArgValAlaArgAlaHisIleuLeuAspHisArgGly 279
 Db 1494 -----GTCCCGCGCGTGAACCGCGGCGCTTACACAGCTGCT 1456
 RESULT 6
 US-08-459-448A-2/C
 ? Sequence 2, Application US/08459448A
 ? Patent No. 5859336
 ? GENERAL INFORMATION:
 ? APPLICANT: Kozietel, Michael G.
 ? APPLICANT: Desai, Nalini M.
 ? APPLICANT: Lewis, Kelly S.
 ? APPLICANT: Kramer, Vance C.
 ? APPLICANT: Warren, Gregory W.
 ? APPLICANT: Eyoia, Stephen V.
 ? APPLICANT: Crossland, Lyde D.
 ? APPLICANT: Wright, Martha S.
 ? APPLICANT: Merlino, Ellis J.
 ? APPLICANT: Launis, Karen L.
 ? APPLICANT: Kolstein, Steven J.
 ? APPLICANT: Bowman, Cindy G.
 ? APPLICANT: Dawson, John L.
 ? APPLICANT: Dunder, Erik M.
 ? APPLICANT: Pace, Gary M.
 ? APPLICANT: Suttie, Jane L.
 ? TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
 ? NUMBER OF SEQUENCES: 94
 ? TITLE OF INVENTION: INSECTICIDAL ACTIVITY IN MAIZE
 ? CORRESPONDENCE ADDRESS:

us-09-709-103-1f4.p2n.rn1

Tue Dec 31 15:27:52 2002

ADDRESSEE: No. 5859336artis Corporation
STREET: Patent & Trademark Dept., 520 White Plains
CITY: Tarrytown
STATE: New York
COUNTRY: USA
ZIP: 10591-9005

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA: US/08/459,448A
FILING DATE: 02-JUN-1995
CLASSIFICATION: 800

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/951,715
FILING DATE: 25-SEP-1992

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/772,027
FILING DATE: 04-OCT-1991

ATTORNEY/AGENT INFORMATION:
NAME: Pace, Gary M.
REGISTRATION NUMBER: 40403

REFERENCE/DOCKET NUMBER: CGC 1577/CTP/DIV4
TELEPHONE: (919)541-8582
TELEFAX: (919)541-8689

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 3468 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "Synthetic DNA"
HYPOTHETICAL: NO
FEATURE:

NAME/KEY: misc feature
LOCATION: 1..3468
OTHER INFORMATION: /product = "Full-length pure maize
optimized synthetic Bt"

OTHER INFORMATION: /note = "Disclosed in Figure 3 as syn1f.mze"
US-08-459-448A-2

Alignment Scores:
Pred. No.: 1.72e-06
Score: 166.00
Percent Similarity: 39.43%
Best Local Similarity: 29.97%
Query Match: 11.76%
DB: 2

US-09-709-103-1f4 (1-282) x US-08-459-448A-2 (1-3468)

QY 1 LeuAlaAspAlaAlaLeuValLeuGly-----LeuAla 13

Db 2304 GTACCGATCAGTAGATCTCCAGGTCCTGCTGTCGATGAGCGCGCAGCTGGTA 2245

QY 14 AlaGlyAlaGlyLeuLeuAlaAspValHisGluValAlaValTyAlaGlyProAlaArg 33

Db 2244 GCGGGTGTAGCGCTTACGTTGCTCTCGATCTCTGTTACAGGTAGGTGGGTAGCA 2185

QY 34 -----GluGlyCysHisAspAlaLysGlyVal----- 42

Db 2184 CTCGTCGAGGTCGCCAGCAGGTCACGTAGTTCTCTTGAACAGCTGTCGCCCGCTG 2125

QY 43 ---AlaArgValAlaAlaAlaAlaAlaAlaGlyProGlnGlnLeuValProGln 61

Db 2124 GATGTGTATGTCGGTGTGCGCGCCAGCGGTCAGCTGCGGTGATGCCCGCGAA 2065

QY 62 ArgLeuLeuValGlnHisValAlaValLeuHisArgAspLeuAlaValGlnValTrpAla 81

Db 2064 GTTGGGGTCTCGCAGCAGGTCGCTCGCTCAGCGCGCTTGGCGTG-----CTT 2014

QY 82 HisLeuAlaGlyGlnLeuGlyHisGlyGluGluArgAlaGluHisLeuValGlnAlaAla 101

Db 2013 CACCTTCTCGCTCAGCTC-----CTTCTTCTCGTCAGCGAGAA 1975

QY 102 ValLeuLeuGlyArgAspLeuGluValGlyAlaAlaLeuGlyValValAlaHisGlnLeu 121

Db 1974 CTC-----GTGCTCAGGCACCTCCACAGGTGCTCAG-----CTG 1939

QY 122 LeuAspLeuAlaLeuValHisLeuAlaValGluValAlaValThrLeuValAlaAlaAsp 141

Db 1938 GTCGATGTGTAGTCGGTCAGCTCGCTCTCAG-----GCCGATCTGTTCTCTCTGT 1885

QY 142 AspGln-----GlyHisValHisValLeuLeuGlyPheValLeuGlu 155

Db 1884 GAACAGCTCGTTCAGCGCTTCTGGCGCGCTCAGGTCTGATCTCGC-----CTCGAA 1831

QY 156 AlaArgLeuGlyValGluAspLeuLeuProGlnProLeuHisLeuLeuGluGlyValAla 175

Db 1830 GGTCACTCGCGCGCACAACTCGATCGCGGTGATGTACACCTCGTTGCCGCTGTGA 1771

QY 176 ValVal-----GlnThrGluHisGlnAspGluAsnValSerCysGluAspGlyGluAla 193

Db 1770 CAGTGGCGCTCAGGTGAACACAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1711

QY 194 ProHisGlyGlyGluArgVal-----ValAlaGlyArgValGlu-----Asp 207

Db 1710 GCCCAGCGTTCGGAAGCTCGCGCTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1651

QY 208 ValGluLeuValAspLeuAlaAla-----AspGlyValGluLeuAlaVal----- 222

Db 1650 GTTCCCTGTTGATGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1591

QY 223 GluValLeuAspGlyArgValGlyValLeuGluAlaAlaGlyGluGluAlaArgHis 242

Db 1590 GCGGTAGCGGATGCGCAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1531

QY 243 AspGlyArgLeuAlaHisLeuGlyArgAlaGluAspAspHisAlaAlaValLeuGly 262

Db 1530 GGTGCTGATCTGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1495

QY 263 ArgAspThrGlnLeuArgValAlaAlaArgAlaHisLeuLeuAspHisArgGly 279

Db 1494 -----GTGCGCGCGCGGTGAAGCGCGCGCGCTTCCACCGCTGCT 1456

RESULT 7

US-08-459-595A-2/c
Sequence 2, Application US/08459595A

Patent No. 6018104
GENERAL INFORMATION:

APPLICANT: Koziel, Michael G.
APPLICANT: Desai, Nalini M.

APPLICANT: Lewis, Kelly S.
APPLICANT: Kramer, Vance C.

APPLICANT: Warren, Gregory W.
APPLICANT: Evola, Stephen V.

APPLICANT: Crossland, Lyle D.
APPLICANT: Wright, Martha S.

APPLICANT: Merlin, Ellis J.
APPLICANT: Launis, Karen L.

APPLICANT: Rothstein, Steven J.
APPLICANT: Bowman, Cindy G.

APPLICANT: Dawson, John L.
APPLICANT: Dunder, Erik M.

APPLICANT: Pace, Gary M.
APPLICANT: Suttie, Janet L.

TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
INSECTICIDAL ACTIVITY IN MAIZE

NUMBER OF SEQUENCES: 94
CORRESPONDENCE ADDRESS:

ADDRESSEE: No. 6018104artis Corporation

Db	2064	GTTCGGGCTCTCCAGACAGAGTTGGCTGTGGCTCCAGCGCTTGCGCTC-----CTT	2014
Qy	82	HTsleuAalagylglnleuGlynhlsylgluIuargAlgluhHlsleuValcIhlaa1a	101
Db	2013	CACCTTCTCGCTCAGCTC-----CTTCTTCTCGCTCAGGCAGAA	1975
Qy	102	ValleuLeuGlyArgAspLeuGluValGlyAlaAlaLeuGlyValValAlahlsIgtleu	121
Db	1974	CTC-----GTGCTCAGGCACCTCCACAGAGTGTCTAC---CTG	1939
Qy	122	LeuAspLeuAlaLeuValHlsleuAlaValGluValAlaValThrlxuValAlaAlaAsp	141
Db	1938	GTTCGATGTGTAGTACGTCTACGTGGCTTCAG-----GCCAGCTGGTGTCTGTCTGT	1885
Qy	142	AspGln-----GlyHlsValHlsValleuLeuGlyPheValleuGlu	155
Db	1884	GAACAGCTCTTCCAGCGCTTCTGGGGGGCGCTCCAGCTGTACTGTGGC-----CTCGAA	1831
Qy	156	AlaArgLeuGlyValGluAspLeuLeuProGluProLeuHlsleuLeuGluGlyValAla	175
Db	1830	GGTCACCTCGGGCGGGACGAACCTGATGGCGTGCATGTACACTGTCTGGCCGCTGTTCAA	1771
Qy	176	ValVal-----GlnHrGluHlsGlnAspGluAsnValSerCyGluAspGlyGluAla	193
Db	1770	CACGGGCGCTCAGGGGTGAACAACGTGCTGCCCTGTGCTGAAGTGAAGGGGGGTGTAA	1711
Qy	194	ProHlsGlyGlyGluArgVal-----ValAlaGlyArgValGlu-----Asp	207
Db	1710	GCCACAGCTGGCGAAGCTGCCCTGTCTGCAGGTGCTGCCCTGCTCATGTGTGGCGCTGAA	1651
Qy	208	ValGluLeuValAspLeuAlaAla-----AspGlyValGluLeuAlaVal-----	222
Db	1650	GTTCCTCGTGTATGGGGCGGGCGCTGCATGCTGTGTGTGAACAGAGTTGGTGTGTCT	1591
Qy	223	GluValLeuAspGlyArgArgValGlyValleuGluAlaAlaGlyGluGluAlaArgHls	242
Db	1550	GGCTTACCGATGCCAGCCGGTAGCGCTGGCTCAGGGGGCGGTGATGTTCACGCCGAG	1531
Qy	243	AspGlyArgLeuAlahlsleuGlyArgAlaGluAspAspHlsAlaIleAlaValleuGly	262
Db	1530	GGTGTGATGTGGCGGGGCTGTGGCGGAGAT-----	1495
Qy	263	ArgAspHrGlnleuArgValAlaArgAlahlsleuLeuAspHlsArgGly	279
Db	1494	-----GTGCGCGCGCGTGAAGCCGGGGCCCTTCACACACGTGT	1456
RESULT 8			
US-08-459-504B-2/c			
Sequence 2, Application US/08459504B			
Patent No. 6075185			
GENERAL INFORMATION:			
APPLICANT: Kozielec, Michael G.			
APPLICANT: Desai, Nalini M.			
APPLICANT: Lewis, Kelly S.			
APPLICANT: Kramer, Vance C.			
APPLICANT: Warren, Gregory W.			
APPLICANT: Ewale, Stephen V.			
APPLICANT: Crossland, Lyle D.			
APPLICANT: Wright, Martha S.			
APPLICANT: Merlin, Ellis J.			
APPLICANT: Lounis, Karen L.			
APPLICANT: Rothstein, Steven J.			
APPLICANT: Bowman, Cindy G.			
APPLICANT: Dawson, John L.			
APPLICANT: Dunder, Erik M.			
APPLICANT: Face, Gary M.			
APPLICANT: Suttie, Janet L.			
TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED			
NUMBER OF INVENTION: INSECTICIDAL ACTIVITY IN MAIZE			
CORRESPONDENCE ADDRESSES: 94			
ADDRESSEE: No. 6075185artis Corporation			
STREET: 3054 Cornwallis Road			

Tue Dec 31 15:27:52 2002

CITY: Research Triangle Park
STATE: NC
COUNTRY: USA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
FILING DATE: 02-JUN-1995
APPLICATION NUMBER: US 08/459,504B
FILING DATE: 25-SEP-1992
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/459,595
FILING DATE: 02-JUN-1995
APPLICATION NUMBER: US 07/951,715
FILING DATE: 25-SEP-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/772,027
FILING DATE: 04-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: CGC1577/CIP/DIV
TELEPHONE: (919)541-8587
TELEFAX: (919)541-8689
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 3468 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "Synthetic DNA"
HYPOTHETICAL: NO
FEATURE:
NAME/KEY: misc feature
LOCATION: 1..3468
OTHER INFORMATION: /product= "Full-length pure maize
OTHER INFORMATION: optimized synthetic Bt"
OTHER INFORMATION: /note= "Disclosed in Figure 3 as synl1t.mze"

US-08-459-504B-2

Alignment Scores:

Pred. No.: 1 72e-06 Length: 3468
Score: 166.00 Matches: 95
Percent Similarity: 39.43% Conservative: 30
Best Local Similarity: 29.97% Mismatches: 120
Query Match: 11.76% Indels: 72
DB: 3 Gaps: 16

US-09-709-103-1f4 (1-282) x US-08-459-504B-2 (1-3468)

QY 1 LeuAlaAspAlaAlaLeuLeuValLeuGly-----LeuAla 13
Db 2304 GTAGCGGATGAGTATCTCCAGGTCTGCTCGATGATGACCGCGAGCTGGTA 2245
QY 14 AlaGlyAlaGlyLeuLeuAlaAspValHisGluValAlaValTyrAlaGlyProAlaArg 33
Db 2244 GCGGGTGTAGGCTTTCAGTGTCTCTCGTATCTTGTATAGGTAGGTGGGTAGCA 2185
QY 34 -----GluGlyCysHisAspAlaLysGlyVal----- 42
Db 2184 CTCGTCGAGGTGCGCCAGCGGGTACGTAGTCTCTTGAACAGCTGCTGCGCGCCCTG 2125
QY 43 ---AlaArgValAlaAlaAlaAlaAlaGlyProGlnGlnLeuLeuValProGln 61
Db 2124 GATGGTATGTCGGTGTGCGCGCCAGCGCGGTCCAGTGGCGGTGATGCGCGGAA 2065
QY 62 ArgLeuLeuValGlnHisValAlaValHisArgAspLeuValGlnValTrpAla 81

Db 2064 GTTGGGTCCTGCAGCAGGTTCGCTCGCTCAGCGCGCTTGGCGTG-----CTT 2014
QY 82 HisLeuAlaGlyGlnLeuGlyHisGlyGluArgAlaGluHisLeuValGlnAlaAla 101
Db 2013 CACCTTCTCGCTCAGCTC-----CTTCTCTGCTCCAGCGCAGAA 1975
QY 102 ValLeuLeuGlyArgAspLeuGluValGlyAlaAlaLeuGlyValValAlaHisGlnLeu 121
Db 1974 CTC-----GTCGCTCAGGCACTCCACAGGTGCTCAGC---CTG 1939
QY 122 LeuAspLeuAlaLeuValHisLeuAlaValGluValAlaValThrLeuValAlaAlaAsp 141
Db 1938 GTCGATGTGTAGTTCGCTCAGCTCGTCTTCAG-----GCCGATCTGTTGCTCTGCT 1885
QY 142 AspGln-----GlyHisValHisValLeuLeuGlyPheValLeuGlu 155
Db 1884 GAACAGCTCGTTACGGCGCTTCTGGCGCGCTCCAGGTGCTACTCGGC-----CTCGAA 1831
QY 156 AlaArgLeuGlyValGluAspLeuProGluProLeuHisLeuGluGlyValAla 175
Db 1830 GGTCACTCGCGCGGCGACCACTCGATGCGGTGATGATGATGATGATGATGATGATGAT 1771
QY 176 ValVal-----GlnThrGluHisGlnAspGluAsnValSerCysGluAspGlyGluAla 193
Db 1770 CACGTGGCGCTCAGGTGAAACACACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1711
QY 194 ProHisGlyGlyGluArgVal-----ValAlaGlyArgValGlu-----Asp 207
Db 1710 GCCACGCTGCGGAAGCTCCCGCTGTCAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1651
QY 208 ValGluLeuValAspLeuAlaAla-----AspGlyValGluLeuAlaVal----- 222
Db 1650 GTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1591
QY 223 GluValLeuAspGlyArgValGlyValGlyValGlyValGlyValGlyValGlyValGly 242
Db 1590 GCGGTAGCGGATGCGACGCGGTAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1531
QY 243 AspGlyArgLeuAlaHisLeuGlyArgAlaGluAspAspHisAlaAlaValLeuGly 262
Db 1530 GGTGCTGATTCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1495
QY 263 ArgAspThrGlnLeuArgValAlaAlaArgAlaHisLeuLeuAspHisArgGly 279
Db 1494 -----GTCGCGCGCGGTGAAAGCGCGCGCGCTTCCACCACGCTGCT 1456

RESULT 9

US-08-459-444-2/c

; Sequence 2, Application US/08459444A

; Patent No. 6121014

; GENERAL INFORMATION:

; APPLICANT: Koziel, Michael G.

; Desai, Nalini M.

; Lewis, Kelly S.

; Kramer, Vance C.

; Warren, Gregory W.

; Evola, Stephen V.

; Crossland, Lyle D.

; Wright, Martha S.

; Merlin, Ellis J.

; Launis, Karen L.

; TITLE OF INVENTION: METHOD FOR PRODUCING A PLANT-OPTIMIZED

; NUCLEIC ACID CODING SEQUENCE

; NUMBER OF SEQUENCES: 94

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: No. 6121014artis Agribusiness Biotechnology Research, Inc.

; STREET: 3054 Cornwallis Road

; CITY: Research Triangle Park

; STATE: NC

; COUNTRY: USA

; ZIP: 27709

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

us-09-709-103-1f4.p2n.rni

Tue Dec 31 15:27:52 2002

QY 208 ValGluLeuValAspLeuAlaAla-----AspGlyValGluLeuAlaVal----- 222
 Db 1650 GTTCCCTGGTGGCGGCGCTCGATGCTGTGGTGGAACTCCAGTTGGTGGTCT 1591
 QY 223 GluValLeuAspGlyArgValGlyValLeuGluAlaAlaGlyGluAlaAlaArgHis 242
 Db 1590 GCGGTAGCGGATCGCACGCGGTAGCGTGGCTCAGGGGGCGGTGATGTTCCACGCGCAG 1531
 QY 243 AspGlyArgLeuAlaHisLeuGlyArgAlaGluAspAspHisAlaAlaValLeuGly 262
 Db 1530 GGTCTGATCTGGCGGCGCTGGTGGCGGCGCAGAT----- 1495
 QY 263 ArgAspThrGlnLeuArgValAlaAlaArgAlaHisLeuLeuAspHisArgGly 279
 Db 1494 -----GTCCGCGCGGTGAAGCCGCGGCGCTTCCACCACTGGT 1456

RESULT 11
 US-09-547-422-2/c
 ; Sequence 2, Application US/09547422
 ; Patent No. 6320100
 ; GENERAL INFORMATION:
 ; APPLICANT: Koziel, Michael G.
 ; Desai, Nalini M.
 ; Lewis, Kelly S.
 ; Kramer, Vance C.
 ; Warren, Gregory W.
 ; Evola, Stephen V.
 ; Crossland, Lyle D.
 ; Wright, Martha S.
 ; Merlin, Ellis J.
 ; Launis, Karen L.

TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
 INSECTICIDAL ACTIVITY IN MAIZE
 NUMBER OF SEQUENCES: 94
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: No. 6320100artis Agribusiness Biotechnology Research, Inc.
 STREET: 3054 Cornwallis Road
 CITY: Research Triangle Park
 STATE: NC
 COUNTRY: USA
 ZIP: 27709

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/547,422
 FILING DATE: 11-Apr-2000
 CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/459,595
 FILING DATE: 02-JUN-1995
 APPLICATION NUMBER: US 07/951,715
 FILING DATE: 25-SEP-1992
 APPLICATION NUMBER: US 07/772,027
 FILING DATE: 04-OCT-1991

ATTORNEY/AGENT INFORMATION:
 NAME: Neigs, J. Timothy
 REGISTRATION NUMBER: 38,241
 REFERENCE/DOCKET NUMBER: S-18805H
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (919)541-8587
 TELEFAX: (919)541-8689

INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3468 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: other nucleic acid
 DESCRIPTION: /desc = "Synthetic DNA"
 HYPOTHETICAL: NO

STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: other nucleic acid
 DESCRIPTION: /desc = "Synthetic DNA"
 HYPOTHETICAL: NO
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: 1..3468
 OTHER INFORMATION: /product= "Full-length pure maize
 OTHER INFORMATION: optimized synthetic Bt"
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 1..3465
 US-09-053-549-3

Alignment Scores:
 Pred. No.: 1.72e-06 Length: 3468
 Score: 166.00 Matches: 95
 Percent Similarity: 39.43% Conservative: 30
 Best Local Similarity: 29.97% Mismatches: 120
 Query Match: 11.76% Indels: 72
 Gaps: 16
 DB:

US-09-709-103-1f4 (1-282) x US-09-053-549-3 (1-3468)

QY 1 LeuAlaAspAlaAlaLeuLeuValLeuGly-----LeuAla 13
 Db 2304 GTACGGATCAGGTAGATCTCCAGGTCCTCGCTGCTCGATGATAGCCGCGAGCTGTA 2245
 QY 14 AlaGlyAlaGlyLeuLeuAspValHisGluValAlaValTyxAlaGlyProAlaArg 33
 Db 2244 GCGGTGTAGCCCTCAGCTGCTCTCGTCGATCTTCTGTCAGGTAGTGGGTAGCA 2185
 QY 34 -----GluGlyCysHisAspAlaLysGlyVal----- 42
 Db 2184 CTCGTGGAAGTGGCCAGCAGGTCACGTAGTCTCTCTGAAACAGTCGTCGCGCCCTG 2125
 QY 43 ---AlaArgValAlaAlaAlaAlaAlaGlyProGlnGlnLeuLeuValProGln 61
 Db 2124 GATGTGATGTCGTGCTGCGCCGCGCCAGCCGCTCAGCTGGCGGTGATGCGCGGAA 2065
 QY 62 ArgLeuLeuValGlnHisValAlaValLeuHisArgAspLeuAlaValGlnValTpAla 81
 Db 2064 GTTGGGCTCTCCAGCAGGTTGCGTCTGCTCAGCGCTTGGCTG-----CTT 2014
 QY 82 HisLeuAlaGlyGlnLeuGlyHisGlyGluGluArgAlaGluHisLeuValGlnAlaAla 101
 Db 2013 CACCTTCTCGCTCAGCTC-----CTTCTCTCTCGTCAGGCAGAA 1975
 QY 102 ValLeuLeuGlyArgAspLeuGluValGlyAlaAlaLeuGlyValValAlaHisGlnLeu 121
 Db 1974 CTC-----GTCCGTAGGCACCTCCAGGTTGCTCAC-----CTG 1939
 QY 122 LeuAspLeuAlaLeuValHisLeuAlaValGluValAlaValThrLeuValAlaAlaAsp 141
 Db 1938 GTCGATGTGTAGTCGTGCTCAGCTGCTTCAG-----GCCATCTGGTTCGTGCT 1885
 QY 142 AspGln-----GlyHisValHisValLeuLeuGlyPheValLeuGlu 155
 Db 1884 GAACAGCTCGTTACGGCTTCTGGGGCGCTCCAGGTCGTACTCGGC-----CTCGAA 1831
 QY 156 AlaArgLeuGlyValGluAspLeuLeuProGlnProLeuHisLeuLeuGlyValAla 175
 Db 1830 GGTACCTCGCGCGGCAGAACTCGATCGCGTGTGATGATCCTGCTGCTGCTGTTGAA 1771
 QY 176 ValVal-----GlnThrGluHisGlnAspGluAsnValserCysGluAspGlyGluAla 193
 Db 1770 CACGTGGCGGTGAGGTGAGACAGCTGCTGCCGCTTCTGAAGTTGAAGGGGTGCTGAA 1711
 QY 194 ProHisGlyGlyGluArgVal-----ValAlaGlyArgValGlu-----Asp 207
 Db 1710 GCCCAGGTGCGAAGCTGCGCTCTGTCAGGTTGCTGCGCGCTGCTGCTGCGGTGAA 1651

```

1      FEATURE:
2      NAME/KEY: misc feature
3      LOCATION: 1..3468
4      OTHER INFORMATION: /product= "Full-length pure maize
5      optimized synthetic Bt"
6      /note= "disclosed in Figure 3 as syn117.mze"
7
8      SEQUENCE DESCRIPTION: SEQ ID NO: 2:
9      US-09-547-422-2

```

Alignment Scores:	
Pred. No.:	1,72e-06
Score:	166.00
Percent Similarity:	39.43
Best Local Similarity:	29.978
Query Match:	11.768
DB:	4
Gaps:	16

US-09-09-103-1F4 (1-282) X US-09-547-422-2 (1-3468)

[illegible]

```

Db      1590 GCGGTAGCGGATCCGACCGCGGTACCGTGCCTAGGGGGCGGTGTTCACCGCGAC 153
Oy      243 AspGlyArgLeuAlaHisLeuGlyArgAlaGluAspAspHisAlaIleAlaValLeuGly 262
Db      1530 GGTGTGATCTGGCGGGGCGTGTGCGCGGAGAT-----1495
Oy      263 ArgAspThrGlnLeuArgValAlaArgAlaHisLeuLeuAspHisArgGly 279
Db      1494 -----GTGCGCGCGGTGAACCGGGGCGCTTTCACCAACCGTGT 1456

RESULT 12
US-07-951-715A-4/c
Sequence 4, Application US/07951715A
Patent No. 5625136
GENERAL INFORMATION:
APPLICANT: Kozielec, Michael G.
APPLICANT: Desai, Nalin M.
APPLICANT: Lewis, Kelly S.
APPLICANT: Kramer, Vance C.
APPLICANT: Warren, Gregory W.
APPLICANT: Evola, Stephen V.
APPLICANT: Crossland, Kyle D.
APPLICANT: Wright, Martha S.
APPLICANT: Merlin, Ellis J.
APPLICANT: Launis, Karen L.
APPLICANT: Rothstein, Steven J.
APPLICANT: Bowman, Cindy G.
APPLICANT: Dawson, John L.
APPLICANT: Dunder, Erik M.
APPLICANT: Pace, Gary M.
APPLICANT: Suttle, Janet L.
TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
NUMBER OF SEQUENCES: 94
CORRESPONDENCE ADDRESS:
ADDRESSEE: CIBA-GEIGY Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: New York
COUNTRY: USA
ZIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30B
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/951,715A
FILING DATE: 25-SEP-1992
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/772,027
FILING DATE: 04-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: Serull, W. Murray
REGISTRATION NUMBER: 32,943
REFERENCE/DOCKET NUMBER: S-18805/A/CGC 1577/CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919)541-8615
TELEFAX: (919)541-8689
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 3468 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "Synthetic DNA"
HYPOTHETICAL: NO
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1..3468
OTHER INFORMATION: /product= "Full length synthetic

```


Db 1530 GGTGCTGATCTGGCCGGGGCTGTGTGGGGCGCAGAT-----1495
 Oy 263 ArgAspThrGlnLeuArgValAlaAlaGlnAhlSLeuLeuAspHisArgGly 279
 Db 1494 -----GTGGCCGCCGGGTGAAGCCGGGGCCCTTCACCAACGCTGT 1456
 RESULT 14
 US-08-459-595A-4/c
 ; Sequence 4, Application US/08459595A
 ; Patent No. 6018104
 GENERAL INFORMATION:
 APPLICANT: Koziel, Michael G.
 APPLICANT: Desai, Nalini M.
 APPLICANT: Lewis, Kelly S.
 APPLICANT: Kramer, Vance C.
 APPLICANT: Warren, Gregory W.
 APPLICANT: Evola, Stephen V.
 APPLICANT: Crossland, Lyle D.
 APPLICANT: Wright, Martha S.
 APPLICANT: Merlin, Ellis J.
 APPLICANT: Launis, Karen L.
 APPLICANT: Rothstein, Steven J.
 APPLICANT: Bowman, Cindy G.
 APPLICANT: Dawson, John L.
 APPLICANT: Dunder, Erik M.
 APPLICANT: Face, Gary M.
 APPLICANT: Suttie, Janet L.
 TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
 NUMBER OF SEQUENCES: 94
 TITLE OF INVENTION: INSECTICIDAL ACTIVITY IN MAIZE
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: No. 6018104artis Corporation
 STREET: Patent & Trademark Dept., 520 White Plains
 STREET: Rd., POB 2005
 CITY: Tarrytown
 STATE: New York
 COUNTRY: USA
 ZIP: 10591-9005
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/459,595A
 FILING DATE: 02-JUN-1995
 CLASSIFICATION: 800
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/951,715
 FILING DATE: 25-SEP-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/772,027
 FILING DATE: 04-OCT-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: Pace, Gary M.
 REGISTRATION NUMBER: 40403
 REFERENCE/DOCKET NUMBER: CGC 1577/CIP/DIV3
 TELEPHONE: (919)541-8582
 TELEFAX: (919)541-8689
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3468 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: other nucleic acid
 DESCRIPTION: /desc = "Synthetic DNA"
 FEATURE:
 HYPOTHEICAL: NO
 NAME/KEY: misc feature
 LOCATION: 1..3468
 OTHER INFORMATION: /product= "Full length synthetic

```

; OTHER INFORMATION: maize optimized"
;
; OTHER INFORMATION: /note="disclosed in Figure 3 as synful.mod. This sequence
; OTHER INFORMATION: identical to fisynt.fin as disclosed in Figure 1."
; OTHER INFORMATION:
US-08-459-595A-4

```

Alignment Scores:	8.06e-06	Length:	3468
Pred. No.:	159.00	Matches:	94
Score:	39.12%	Conservative:	30
Percent Similarity:	39.12%	Mismatches:	121
Best Local Similarity:	29.65%	Indels:	72
Query Match:	11.27%	Gaps:	16
DB:	3		
US-09-709-103-1F4 (1-282) x US-08-459-595A-4 (1-3468)			
Qy	1	LeuAlaAspAlaAlaLeuValLeuGly	-----LeuAla 13
Db	2304	GTACCGGATCAGGTAGATCTTCAGGTCTGGCTGCTCGATGTAGCCGCGAGCTGGTA	2245
Qy	14	AlaGlyAlaGlyLeuLeuAlaAspValHisGluValAlaValTrpAlaArg 33	
Db	2244	GCSSGTGTAGGCTTCAGGTGCTCTCGATCTTCGTACAGGTGGGTAGCA	2185
Qy	34	-----GluGlyCysHisAspAlaGlyVal	----- 42
Db	2184	CTGCTCGAAGTCCCGCAGCAGGTCCAGTAGTCTCTTGAACACGTCGTCGCCCGCTG	2125
Qy	43	---AlaArgValAlaAlaAlaAlaAlaAlaGlyProGlnGlnLeuValProGln 61	
Db	2124	GATGGTATCTCGGTCTCCCGCGCAGCCGCGTCCAGTGGCGGTGATGCCGCGAA	2065
Qy	62	ArgLeuLeuValGlnHisValAlaValLeuHisArgAspLeuAlaValGlnValTrpAla 81	
Db	2064	GTTGGGTCTCGCAGCAGGTTCGGCTCGTCCAGCGCTTGGCGTG	2014
Qy	82	HisLeuAlaGlyGlnLeuGlyHisGlyGluGluArgAlaGluHisLeuValGlnAlaAla 101	
Db	2013	CACCTCTCGCTCAGTCTC-----CTTCTTCGCTCCAGGCAGAA	1975
Qy	102	ValLeuLeuGlyArgAspLeuGluValGlyAlaAlaLeuGlyValValAlaHisGlnLeu 121	
Db	1974	CTC-----GTCGCTCAGGCACCTCCACCAGGTGCTCAC---CTG	1939
Qy	122	LeuAspLeuAlaLeuValHisLeuAlaValGluValAlaValThrLeuValAlaAlaAsp 141	
Db	1938	ATCGATGTGTAGTCCGTCAGCTCCAGCTTCAG-----GCCGATCTGGTGTCTGCTGCT 1885	
Qy	142	AspGln-----GlyHisValHisValLeuLeuGlyPheValLeuGlu 155	
Db	1884	GAACAGCTCGTTACGGCTTCGTAGCCCTCTCCAGCTCTCCAGGTCGTAACGCGC-----CTCGAA	1831
Qy	156	AlaArgLeuGlyValGluAspLeuLeuProLeuHisLeuLeuGluGlyValAla 175	
Db	1830	GGTCACTCCGGCGGCACGAATCGATCGGTGATGATACCTCTGTCGGCTGTGAA	1771
Qy	176	ValVal-----GlnThrGluHisGlnAspGluAsnValSerCysGluAspGlyGluAla 193	
Db	1770	CACGTGGCGCTCAGGTGAACACGCTGCTCGCTGCTGAAGTTGAAGGGGGTGGTGA	1711
Qy	194	ProHisGlyGlyGluArgVal-----ValAlaGlyArgValGlu-----Asp 207	
Db	1710	GCCACGGTGGGAAGTCCGCGCTCTCAGGTTCCTGCCCTGCTCATGGTGGCGCTGAA	1651
Qy	208	ValGluLeuValAspLeuAlaAla-----AspGlyValGluLeuAlaVal----- 222	
Db	1650	GTTGCCCTGTTGATGGGGCGCCGCTCGATGCTGGTGGAACTCGACGTGGTGGTCT	1591
Qy	223	GluValLeuAspGlyArgArgValGlyValLeuGluAlaAlaGlyCylGluAlaArgHis 242	
Db	1590	GGCGTACCGGATCCGAGCCGCTGCGTGGCTCAGGGGGCGGTGATGTTCCAGCGCAG	1531
Qy	243	AspGlyArgLeuAlaHisLeuGlyArgAlaGlyAspHisAlaLeuAlaValLeuGly 262	

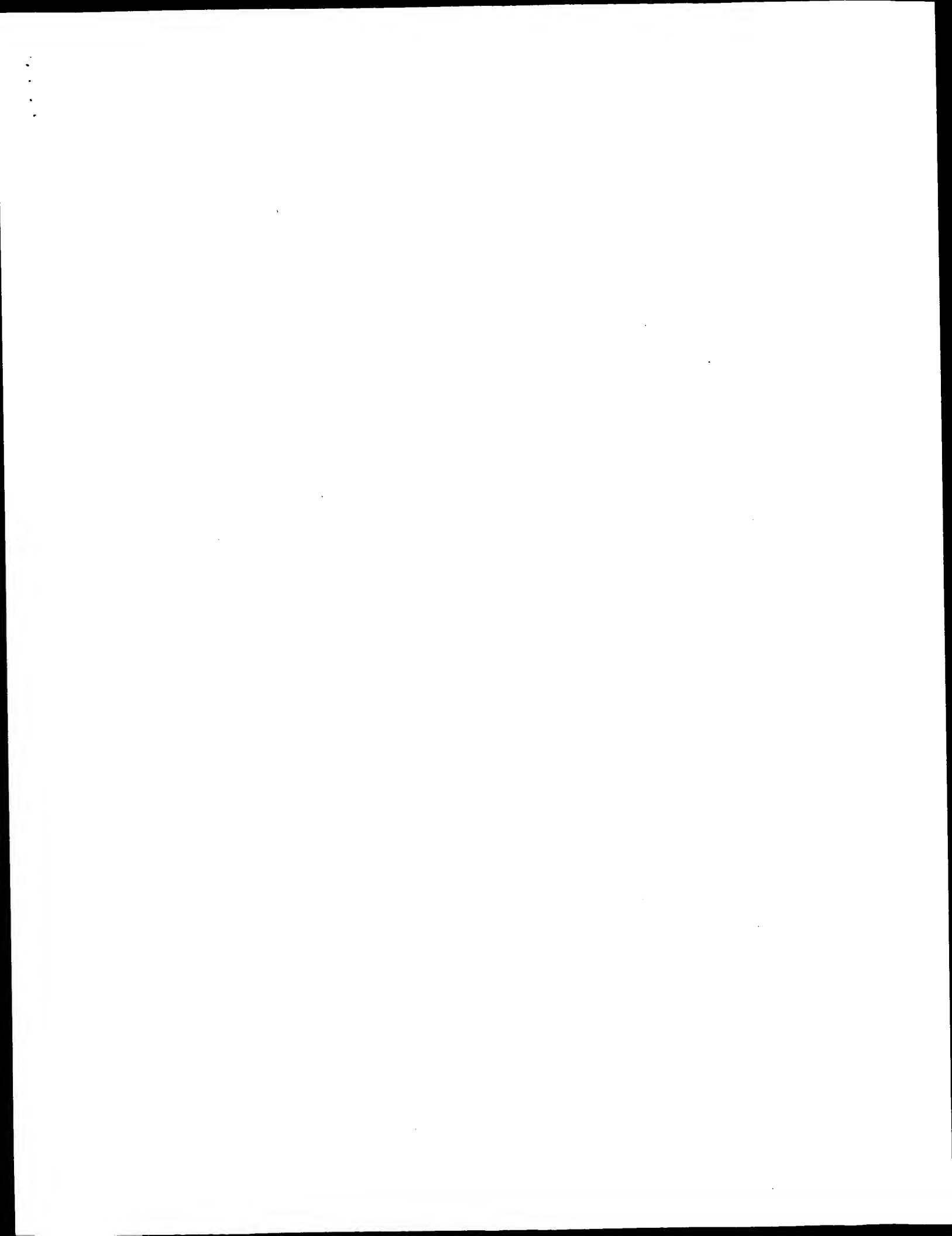
-----TCTGGGGCAGGAT----- 1495

D8	1530	GATGCTGAACAGCGGGGCCTTGTCCGCCCGCAGA..
Qy	263	ArgAspThrClnLeuArgValAlaAargAlahisLeuLeuAspHis=ArgGly
Dh	1494	::::-----GTGCGCCCGGTGAACCGGGGCCCTTTACCACGCCTGGT

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999

```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on:

December 30, 2002, 15:42:56 ; Search time 2179 Seconds
(without alignments)
2095.974 Million cell updates/sec

Title: US-09-709-103-1f4
Perfect score: 1411
Sequence: 1 LADDAALVVGIAAGAGLAA.....RDTQLVARAHLLDRGQPF 282

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlp
-O/cgnt_1/USPTO.spool/SUILLIVANI4/runat_30122002_144047_21512/app.query.fasta_1.455
-DB=EST -OFAST=fastp -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOCCALIGN=200 -THR_SCORE=Pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTPRT=pco -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=SULLIVANI4@CGN 1 1 1716 @runat_30122002_144047_21512 -NCPUS=6 -ICPU=3
-NO_XLPHY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estcom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pin:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_ther:*
26: em_gss_pro:*
27: em_gss_tod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1194	84.6	1035	14	BM919341
2	1131.5	80.2	962	13	BM543472
3	1112	78.8	1103	14	BM920514
4	1111	78.7	742	10	AM028127
5	1098	77.8	965	9	AL533318
6	1082	76.7	958	14	BM719566
7	1044	74.0	1053	14	BM921737
8	1029	72.9	1032	14	BM067637
9	1005	71.2	1023	13	BM543630
10	896	63.5	758	13	BM553776
11	894.5	63.4	814	13	BM596688
12	893	63.3	1160	14	BM805574
13	881	62.4	1137	14	BM921656
14	859.5	60.9	831	13	BM596637
15	832	59.0	506	13	BM311047
16	808	57.3	699	12	BM706012
17	792	56.1	699	12	BM709229
18	792	56.1	700	13	BM596509
19	792	54.6	655	12	BM711792
20	757	52.2	674	13	BM101563
21	737	52.2	666	13	BM549939
22	729	51.7	675	13	BM545172
23	716	50.7	649	12	BM085090
24	659	48.7	617	13	BM333669
25	659	48.7	617	13	BM333669
26	619	43.9	571	13	BM426066
27	618	43.8	589	12	BM682922
28	614.5	43.6	726	12	BM426066
29	601.5	42.6	1404	13	BM548645
30	599	42.5	556	12	BM654362
31	590	41.8	644	13	BM969048
32	587.5	41.6	644	13	BM912923
33	580	41.1	648	10	BM636889
34	576	40.8	689	10	BM632699
35	575	40.8	531	13	BM823596
36	575	40.8	540	13	BM824314
37	575	40.8	541	13	BM759201
38	575	40.8	556	13	BM1913105
39	575	40.8	1006	14	BM073742
40	571	40.5	947	14	BM954076
41	570	40.4	539	13	BM760810
42	568	40.3	520	13	BM823533
43	562	39.8	535	13	BM763426
44	548	38.7	539	13	BM191057
45	546	38.7	1332	13	BM460899

ALIGNMENTS

RESULT 1
BM919341/c
LOCUS
DEFINITION
AGENCYCOURT_6715681 NIH_MGC_120 Homo sapiens CDNA clone IMAGE:5748585
5', mRNA sequence.
ACCESSION
BM919341
VERSION
BM919341.1 GI:19369720
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 1035)
NIH-MGC <http://mgi.nci.nih.gov/>.
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL
COMMENT

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabps-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.N.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found at: <http://www.ncbi.nlm.nih.gov/Genbank/LLNL> at:

```

found through the IIR.A.G.R.: Compendium
http://image.llnl.gov
Plate: LLAM12776 row: j column: 10
High quality sequence stop: 638.
Location/Qualifiers
    1..1035
       /organism="Homo sapiens"
       /db_xref="taxon:9606"
       /clone="IMAGE:5748585"
       /clone_lib="NIH MGC_120"
       /lab_host="DH10B"
       /note="Organ: pooled pancreas and spleen; Vector:
pcmv-SpOrf6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of spleen and pancreas from 28 y
male. Library is oligo-dr primed and directionally cloned
(EcoRV site is destroyed upon cloning). Average insert
size 1.5 kb, insert size range 1-2.5 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (Invitrogen). Research Genetics
tracking code 025. Note: This is a NIH MGC Library."
          293 g      146 t

```

QY	174	ValAlaValValGlnThrGluHisGlnAspGluValSerCysGluAspGlyGluAla	193
Db	513	GTTCGGGTGTTCAGACTGACACACGAGTAAAAACGTCTCTCTGAGGATGAGAGCGC	454
QY	194	ProHisGlyGlyGluArgValAlaAlaGlyArgValGluAspValGluLeuValAspLeu	213
Db	453	CCGCATGGCGGGAACGGGTGGTTGTCGGAGCGTGTCTGAGGATGTCAGCTGGTAGACCTC	394
QY	214	AlaAlaAspGlyValGluLeuAlaValGluValLeuAspGlyArgArgValGlyValLeu	233
Db	393	GCCGCGGATGGAGTAGAACTTCGGGTGAAGTCCTCGATGTTAGGCGTGTAGGCTCCTC	334
QY	234	GluAlaAlaGlyGluAlaAlaArgHisAspGlyArgLeuAlaHisLeuGlyArgAlaGlu	253
Db	333	GAAGCGCGGTGAGGAGCGGACACGATGCCGTCTTGCCACCTTGGACGAGCCGAG	274
QY	254	AspAspHisAlaIleAlaValLeuGlyArgAspThrGlnLeuArgValAlaArgAlaHis	273
Db	273	GATGACCATCGATGACGAGTCTTTGCGCCGGGATACTCAGTCCGAGTCTCGGCGACAT	214
QY	274	LeuLeuAspHisArgGlyGlnPheHis	282
Db	213	CTTCTTGATCATCGGCCCATGTTTCAT	187
RESULT 2			
BM543472/c		962 bp	mRNA linear EST 20-FEB-2002
LOCUS	BM543472	AGCENQ027 6492614	NIH_MGC_124 Homo sapiens cDNA clone IMAGE:5726554
DEFINITION		5', mRNA sequence.	
ACCESSION	BM543472		
VERSION	BM543472.1	GI:18773895	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
	1 (bases 1 to 962)		
REFERENCE	NIH-MGC http://mgc.nci.nih.gov/ .		
AUTHORS	National Institutes of Health, Ph.D.		
TITLE	Unpublished (1999)		
JOURNAL	Contact: Robert Strausberg, Ph.D.		
COMMENT	Email: cgapbs-remail.nih.gov		
	Tissue Procurement: Invitrogen		
	cDNA Library Preparation: Life Technologies, Inc.		
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)		
	DNA Sequencing by: Agencourt Bioscience Corporation		
	DNA Distribution: MGC clone distribution information can be		
	Clone Distribution: www.igep.org/ Consortium/LLNL at:		

FEATURES source

```

1. .962
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5726554"
/clone_lib="NIH_MGC_124"
/tissue_type="hippocampus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMV-SPORT6;
(destroyed); Site 2; NotI; RNA source male;
age 27. Library is oligo-dT primed and dir-
cloned (EcoRV site is destroyed upon clone-
insert size 1.4 kb, insert size range 0.9-
normalized and enriched for full-length cl-
constructed by C. Gruber (Invitrogen). Re-
tracking code 012."
acc3 272 141 t 1 others

```


us-09-709-103-1f4.p2n.rst

Tue Dec 31 15:27:53 2002

```

Db 612 CGTCCAGTCTCTCTGTTTGTCTTGGCAAGACTTGGTGTGAGGATCTGCTGCC 553
Qy 165 oGluProLeuHisLeuGluGlyValAlaValAlaValGlnThrGluHisGlnAspGluAs 185
Db 552 TGAGCCGCTGCACCTCTCGAAGGAGTACCGTTGTCAGACTGACACACAGGATGAAAA 493
Qy 185 nValSerCysGluAspGlyGluAlaProHisGlyGlyGluArgValAlaGlyArgVa 205
Db 492 GGTCTCTGAGGATGAGAGGCGCCGATGCGCGGAAACGGGTGTTGCGGAGCTGT 433
Qy 205 lGluAspValGluLeuValAspLeuAlaAspGlyValGluLeuAlaValGluValLe 225
Db 432 CGAGGATGTCAGCTGTAGTACCTCGCGCGGATGAGTGAAGTGAAGTCTCT 373
Qy 225 uAspGlyArgValGlyValLeuGluAlaAlaGlyGluGluAlaAArgHisAspGlyAr 245
Db 372 CGATGTTAGCGTGTAGCGTCTCGAAGCGCGGTGAGGAAGCGGACACAGATGCCG 313
Qy 245 lGluAlaHisLeuGlyArgAlaGluAspAspHisAlaAlaValLeuGlyArgAspTh 265
Db 312 TCTTGCACACCTTGGACGAGCGAGGATGACATCGGATACAGTCTTGGCGGATAC 253
Qy 265 rGlnLeuArgValAlaArgAlaHisLeuAspHisArgGlyGlnPheHis 282
Db 252 TCAGCTCCGAGTGTCTCGGCGACATCTTCTTGATCATCGCGCGCAGTTTCAT 201

RESULT 4
AW028127/c 742 bp mRNA linear EST 27-OCT-1999
LOCUS wv26c07.x1 NCI CGAP Kid11 Homo sapiens cDNA clone IMAGE:2530668 3'
DEFINITION similar to TR_035626 035626 RAS, DEXAMETHASONE-INDUCED 1, mRNA
sequence.
ACCESSION AW028127
VERSION AW028127.1 GI:5866883
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 742)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLES National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
CDNA Sequencing by: Washington University Genome Sequencing Center
Clone Distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bbrp/image/image.html
Seq primer: -400P from Gibco
High quality sequence stop: 445.
Location/Qualifiers
1. 742
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2530668"
/clone_lib="NCI CGAP_Kid11"
/lab_host="pH105"
/note="Organ: kidney; Vector: pTVT3D-Pac (Pharmacia) with
a modified polylinker; Site 1: Not 1; Site 2: Eco RI;
plasmid DNA from the normalized library NCI CGAP_Kid3 was
prepared, and ss circles were made in vitro. Following HAP
purification, this DNA was used as tracer in a subtractive
hybridization reaction. The driver was PCR-amplified cDNAs
from a pool of 5,000 clones made from the same library
(clone109 1322376-1323911, 1456007-1456775, and
1500552-1502855). Subtraction by Bento Soares and M.

```

```

BASE COUNT 149 a 260 c 208 g 122 t 3 others
ORIGIN
Alignment Scores: 3.72e-97 Length: 742
Pred. No.: 1111.00 Matches: 225
Score: 96.19% Conservative: 2
Percent Similarity: 95.34% Mismatches: 9
Best Local Similarity: 78.74% Indels: 1
Query Match: 10 Gaps: 0
DB:
US-09-709-103-1f4 (1-282) x AW028127 (1-742)
Qy 47 AlaAlaAlaAlaAlaGlyProGlnGlnLeuValProGlnArgLeuValGln 66
Db 741 GCACGACCCGCGCTGC-GGACCGCACAGGCTTCTGTTTCGCGACGGCTTCTGTGCGAGG 683
Qy 67 HisValAlaValLeuHisArgAspLeuAlaValGlnValTrpAlaHisLeuAlaGlyGln 86
Db 682 CACGTCAGGTACTGACACGAGACCTTGGCGTGCAGGTCTGGGCTCATCTCGCTGGGCAG 623
Qy 87 LeuGlyHisGlyGluArgAlaGluHisLeuValGlnAlaAlaValLeuLeuGlyArg 106
Db 622 CTTGGCCATGCGAAGAGCGCGGACACTCTGCTCGAGCTGCTGTTCTTCTTGGCNGA 563
Qy 107 AspLeuGluValGlyAlaAlaLeuGlyValAlaHisGlnLeuLeuAspLeuAlaLeu 126
Db 562 GATCTCGAAGTAGGCGGACGCGTGGGGTCTGCGCCACGAGCTGCTCGATCTCGCGCTG 503
Qy 127 ValHisLeuAlaValGluValAlaValThrLeuValAlaAlaAlaAspAspGlnGlyHisVal 146
Db 502 GTCCACCTCGCGGTAGAGTCCGCTTGTCCAGACTGAACACACAGGATGAACAGTC 443
Qy 147 HisValLeuLeuGlyPheValLeuGluAlaArgLeuGlyValGluAspLeuLeuProGlu 166
Db 442 CACGTCTCTTGGTTTGTCTTGGGACAGACTTGGTGTCCAGGATCTGCTGCTGAG 383
Qy 167 ProLeuHisLeuLeuGluGlyValAlaValValGlnThrGluHisGlnAspGluAsnVal 186
Db 382 CCGCTGCACCTCTCGAAGGAGTCCGCTTGTCCAGACTGAACACACAGGATGAACAGTC 323
Qy 187 SerCysGluAspGlyGluAlaProHisGlyGlyGluArgValAlaAlaGlyArgValGlu 206
Db 322 TCTGTGAGGTAGAGAGCGCGCATGCGGGGAAACGGTGTGCGCGAGCTGTCGAG 263
Qy 207 AspValGluLeuValAspLeuAlaAlaAspGlyValGluLeuAlaValGluValLeuAsp 226
Db 262 GATGTCGAGCTGTGATGACCTCGCGCGGATGGAGTAGAACTTTCGCGTGAAGTCTCTCAT 203
Qy 227 GlyArgArgValGlyValLeuGluAlaAlaGlyGluAlaAlaArgHisAspGlyArgLeu 246
Db 202 GGTAGGCGTGTAGCGTCTTCCGAGCGCGAGGTGAGGAAGCGCGACGATGGCGCTCTT 143
Qy 247 AlaHisLeuGlyArgAlaGluAspAspHisAlaAlaValLeuGlyArgAspThrGln 266
Db 142 GCCCACCTTCGACGAGCGCGAGGATGACCATCGATAGCAGTTCTTGGCGCGGATACTCAG 83
Qy 267 LeuArgValAlaAlaArgAlaHisLeuLeuAspHisArgGlyGlnPheHis 282
Db 82 CTCGAGTCTGCTCGGCGACATCTTCTTGATCATCGCGCGCAGTTTCAT 35

RESULT 5
AL533318/c
LOCUS AL533318 LTI FL015 Brnl Homo sapiens cDNA clone CS0DN003VJ19 5
DEFINITION prime, mRNA sequence.
ACCESSION AL533318
VERSION AL533318.1 GI:12796811
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

REFERENCE 1 (bases 1 to 965)
 AUTHORS Li, W. B., Gruber, C., Jesse, J. and Polyses, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: seque@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
 source
 1...965
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CS0DN0031019"
 /clone_lib="LTI_Fl015_Brn1"
 /sex="male"
 /tissue_type="Adult brain"
 /note="Vector: PCMVSPORT 6; Site 1: NotI, 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the PCMVSPORT 6 vector. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 189 a 352 c 273 g 144 t 7 others

Alignment Scores:
 Pred. No.: 9,65e-96 Length: 965
 Score: 1098.00 Matches: 231
 Percent Similarity: 92.40% Conservative: 0
 Best Local Similarity: 92.40% Mismatches: 17
 DB: 77.82% Indels: 4
 Gaps: 1

US-09-709-103-1f4 (1-282) x AL533318 (1-965)

QY 33 ArgGlyGlyCyHAspAlaLysGlyValAlaArgValAlaAlaAlaAlaAla 52
 Db 932 CGCGRAGGTGCACGATGCAAAA-----GCTGCCCGGCTGCCCGCGCGCTG 879
 QY 53 GYProGlnGlnLeuLeuValProGlnArgLeuLeuValGlnHisValAlaValLeuHis 72
 Db 878 CGCGCCGACGACGATTTTGTTCGCGAGGCTTCTTCTGTCGACGACCTCCGACGACGAC 819
 QY 73 ArgAspLeuAlaValAlaGlnValTTPAlaHisLeuAlaGlyGlnLeuGlyHisGlyGln 92
 Db 818 CGAGACTT-GCGGTGCGAGGTCTGCGCTCATCTGCTGCGGAGCTTGGCCATGCGCAGAG 760
 QY 93 ArgAlaGlnHisLeuValGlnAlaAlaValLeuLeuGlyArgAspLeuGlyValGlyVal 112
 Db 759 CGCGCGACACATCTGCTGCTGCGAGGCTGCTCTTCTTCTGCGCGAGATCTCGAATGAGCGCA 700
 QY 113 AlaLeuGlyValAlaAlaHisGlnLeuLeuAspLeuAlaLeuValHisLeuAlaValGly 132
 Db 699 GCGCTGGGGTCTCTGCGCCACGAGCTGCTGATCTGCGCTGCTGCTGCTGCTGCTGCTGCTG 640
 QY 133 ValAlaValThrLeuValAlaAlaAspAspGlnGlyHisValHisValLeuLeuGlyPhe 152
 Db 639 GTGGCGGTCTCCCTTGTGCGCGAGATGACGAGGCGACGTCACGTTCTCTCTGTTT 580
 QY 153 ValLeuGlnAlaArgLeuGlyValGlnAspLeuLeuProGlnProLeuHisLeuLeuGly 172
 Db 579 GTTCTTGAAGGACAGACTTGTGTGCGAGATCTGCTGCTGAGCGCTGCACTCTCTGAA 520
 QY 173 GlyValAlaValAlaGlnThrGlnHisGlnAspGlnAspValSerCysGlnAspGlyGln 192
 Db 519 GAGGTGCGGTGTCTGACGACTGAACACGAGATGAAAAAGTCTCTGTGAGATGAGAG 460
 QY 193 AlaProHisGlyGlyGlnArgValAlaAlaGlyArgValAlaGlnAspValGlnLeuValAsp 212

Db 459 CGCGC-CATGCGCGGGAACGGGTGTGTCGCGGACGCTGACGATGTGAGTAC 401
 QY 213 LeuAlaAlaAspGlyValGlnLeuAlaValGlnValLeuAspGlyArgArgValGlyVal 232
 Db 400 CTCGCCCGGATGATGATGAACTTGGCGGTGAAGTCTTCGATGTGAGGCGTGTAGGGCTC 341
 QY 233 LeuGlnAlaAlaGlyGlnGlnAlaArgHisAspGlyArgLeuAlaHisLeuGlyArgAla 252
 Db 340 CTGAAGCGCGCGGTGAGGAGACGACGACATGCGCTCTTCCCACTTGACAGAGCC 281
 QY 253 GluAspAspHisAlaLeuAlaValLeuGlyArgAspThrGlnLeuArgValAlaArgAla 272
 Db 280 GAGGATGACCAATGCCATGACAGTCTTGGCGGAGATCTGACGTCCGAGTCTCGGGA 221
 QY 273 HisLeuLeuAspHisArgGlyGlnPheHis 282
 Db 220 CATCTTGTGATCATGCGCGCACTTCAT 191

RESULT 6
 BQ719566/c 958 bp mRNA linear EST 16-JUL-2002
 LOCUS AGENCOURT 8305191 lukeki_sympathetic_trunk Homo sapiens cDNA clone
 DEFINITION IMAGE:6193170 5', mRNA sequence.
 ACCESSION BQ719566 GI:21858463
 VERSION BQ719566
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 958)
 AUTHORS NIH-MGC http://mgi.cni.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Dr. James R. Lupski
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINTL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LINTL at:
 http://image.lnl.gov
 Plate: L1AM1356 Row: f Column: 19
 High quality sequence stop: 514.
 Location/Qualifiers
 1..958
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:6193170"
 /clone_lib="Lupski_sympathetic_trunk"
 /sex="male"
 /tissue_type="sympathetic trunk"
 /dev_stage="adult, 16 yr"
 /lab_host="DH10B"
 /note="Vector: PCMV-SPORT6 (Life Technologies); Site 1:
 NotI; Site 2: SalI; cDNA made by oligo-dT priming.
 Directionally cloned using the following adaptors:
 5'-TCGACCCAGCGCTCCG-3' and
 1 kb for average insert length 1.9 kb. This is a primary
 library, non-amplified. Library constructed by Life
 Technologies and donated by J. Lupski, M.D., Ph.D. (Baylor
 College of Medicine); available through Life
 Technologies."

FEATURES

source

BASE COUNT 185 a 339 c 275 g 159 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 3.33e-94 Length: 958
 Score: 1082.00 Matches: 233
 Percent Similarity: 83.75% Conservative: 4


```

5', mRNA sequence.
BM543630
VERSION BM543630.1 GI:18774186
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1023)
NIH-MGC http://mgi.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgabs-r@mail.nih.gov
Tissue Procurement: Invitrogen
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM12719 row: m column: 05
High quality sequence stop: 637.
FEATURES
Location/Qualifiers
1..1023
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5726764"
/clone_lib="NIH_MGC_124"
/tissue_type="hippocampus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMV-SPORT6; Site 1: EcoRV
(destroyed); Site 2: NotI: RNA source male hippocampus,
age 27. Library is oligo-dT primed and directionally
cloned (EcoRV site is destroyed upon cloning). Average
insert size 1.4 kb, insert size range 0.9-4 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (Invitrogen). Research Genetics
tracking code 012."
BASE COUNT 196 a 378 c 281 g 168 t
ORIGIN
Alignment Scores:
Pred. No.: 9,6e-87 Length: 1023
Score: 1005.00 Matches: 206
Percent Similarity: 98.10% Conservative: 1
Best Local Similarity: 97.63% Mismatches: 3
Query Match: 71.23% Indels: 3
DB: 13 Gaps: 0
US-09-709-103-1f4 (1-282) x BM543630 (1-1023)
QY 75 LeuAlaValGlnValTrpAlaHisLeuAla--GlyGlnLeuGlyHisGlyGluAlaArg 94
|||||
Db 775 TTGCGGTGCGAGGTCCTGGGCTCATCTCGCGGGCGAGCTTGGCCATGGAAAGAGCGCG 716
|||||
QY 94 laGluHisLeuVal-GlnAlaAlaValLeuLeuGlyArgAspLeuGluValGlyAla 113
|||||
Db 715 CGGAACATCTGGTCCCGAGGTCGTCTTCTTGGCGGAGATCTCGAAGTAGCGGAGCG 656
|||||
QY 114 LeuGlyValValAlaHisGlnLeuLeuAspLeuAlaLeuValHisLeuAlaValGluVal 133
|||||
Db 655 CTGGGGGTGCGCGCCACCGAGCTCTCGATCTCGCGTGTCTCACCTCGCGGTAGAGTC 596
|||||
QY 134 AlaValThrLeuValAlaAlaAspAspGlnGlyHisValHisValLeuLeuGlyPheVal 153
|||||
Db 595 GCGGTACACCTTGTTCGCCGAGATGACACGAGGCGACCTCCACGCTTCTCTGGTTTGT 536
|||||
QY 154 LeuGluAlaArgLeuGlyValGluAspLeuLeuProGluProLeuHisLeuLeuGly 173
|||||
Db 535 CTTGAGGCAAGACTTGGGTGTCGAGATCTGCTGCTGAGCGGCTGCACCTCTCTCGAAGA 476
|||||
QY 174 ValAlaValGlnThrGluHisGlnAspGluAsnValSerCysGluAspGlyGluAla 193
|||||
Db 475 GTCCGGTGTCCAGACTGAACACAGGATGAAACAGTCTCTGTGAGGATGGAGGCG 416
|||||
QY 194 ProHisGlyGlyGluArgValValAlaGlyArgValGluAspValGluLeuValAspLeu 213
|||||
Db 415 CCGCATGGCGGGAACGGGTGGTTCGCCGAGCTGTCGAGGATGTCGAGCTGGTAGACCTC 356
|||||
QY 214 AlaAlaAspGlyValGluLeuAlaValGluValLeuAspGlyArgArgValGlyValLeu 233
|||||
Db 355 GCCCGGATGAGTAGTAACCTTCGGTGGAAAGTCTCTCGATGGTAGCGGTAGGCGTCTC 296
|||||
QY 234 GluAlaAlaGlyGluGluAlaArgHisAspGlyArgLeuAlaHisLeuGlyArgAlaGlu 253
|||||
Db 295 GAAGCGCGCGGTAGGAAGCGGACACGATGGCCGCTCTTCCACCTTGGACGAGCGAG 236
|||||
QY 254 AspAspHisAlaAlaAlaValLeuGlyArgAspThrGlnLeuArgValAlaAArgAlaHis 273
|||||
Db 235 GATGACCATGCGATAGCAGTCTTGGCCGGGATCTCAGCTCCGAGTCCGTCGGGCACAT 176
|||||
QY 274 LeuLeuAspHisArgGlyGlnPheHis 282
|||||
Db 175 CTTCTTGCATCATCGCGGCCAGTTTCAT 149
|||||
RESULT 10
BI553776/c 758 bp mRNA linear EST 05-SEP-2001
LOCUS 603190722F1 NIH_MGC_95 Homo sapiens cDNA clone IMAGE:5262101 5',
DEFINITION mRNA sequence.
ACCESSION BI553776
VERSION BI553776.1 GI:15441088
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 758)
NIH-MGC http://mgi.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgabs-r@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM11660 row: 1 column: 06
High quality sequence stop: 738.
FEATURES
Location/Qualifiers
1..758
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5262101"
/clone_lib="NIH_MGC_95"
/tissue_type="hippocampus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescriptR (modified
pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gtcggg
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',
size-selected for average insert size 2.5 kb and
normalized to 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT 155 a 252 c 249 g 102 t
ORIGIN
Alignment Scores:
Pred. No.: 2e-76 Length: 758

```


Tue Dec 31 15:27:53 2002

```

RESULT 12
BM05574/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source
BASE COUNT 207 a 473 c 279 g 201 t
ORIGIN
Alignment Scores:
Pred. No.: 7,12e-76 Length: 1160
Score: 893.00 Matches: 208
Percent Similarity: 76.49% Conservative: 10
Best Local Similarity: 72.98% Mismatches: 49
Query Match: 63.29% Indels: 18
DB: 14 Gaps: 6
US-09-709-103-1f4 (1-282) x BM05574 (1-1160)
QY 8 LeuValLeuGlyLeuAlaAlaGlyAlaGlyLeuLeuAlaAspValHisGluValAlaVal 27
Db 1019 CTTTGGTTAGGAGGTACCGGGGGGTGGTGG-----GTT 984
QY 28 TyraLaGlyProAlaArgGluGly----CysHisAspAlaGlyValAlaArg----- 44
Db 983 ACCGTGGGGCGGAAAGGAGGTGGCGGGAATCCAGAGGGGGCGCGGGGTAA 924
QY 45 ValAlaAlaAlaAlaAlaAlaAlaGlyProGlnGlnLeuLeu---ValProGlnArgLeu 63
Db 923 ACACCAACCAACCGCGGGGGCGGCGGCAAGCAACTGTGGGTTTCAGCAGGGCCTTT 864
QY 64 LeuValGlnHisVal-AlaVal---LeuHisArgAspLeuAlaValGlnValTrpAlaHi 82
Db 863 TTGGTCAGCATGTTGGCAGTGGGGCGCGGAGCCCTCGCGGGGCGAGGTTCTGGGGTCA 804

```

BM05574 1160 bp mRNA linear EST 05-MAR-2002
 AGENCOURT_6490388 NIH_MGC_124 Homo sapiens cDNA clone IMAGE:5729129
 5', mRNA sequence.
 BM05574 1 GI:19122397
 EST.
 human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1160)
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Invitrogen
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM12725 row: 0 column: 18
 High quality sequence start: 23
 High quality sequence stop: 501.
 Location/Qualifiers
 1..1160
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5729129"
 /clone_lib="NIH_MGC_124"
 /tissue_type="hippocampus"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: EcoRV
 (destroyed); Site_2: NotI; RNA source male hippocampus,
 age 27. Library is oligo-dT primed and directionally
 cloned (EcoRV site is destroyed upon cloning). Average
 insert size 1.4 kb, insert size range 0.9-4 kb. Library is
 normalized and enriched for full-length clones and was
 constructed by C. Gruber (Invitrogen). Research Genetics
 tracking code 012."
 207 a 473 c 279 g 201 t

```

QY 82 sleuAla---GlyGlnLeuGlyHisGlyGluGluArg-AlaGluHisLeuVal-GlnAla 100
Db 803 TCTCCGCTGGGGCCAGCTTGACCAATGACAAAGAGCGCGGGGAGCATGGGTTCCAGGT 744
QY 101 AlaValLeuLeuGlyArgAspLeuGluValGlyAlaAlaLeuGlyValValAlaHisGln 120
Db 743 GGTGTTTTCCTGGGCGGAGATCTCGNAGTAGGCGACGCTGGGGGTCGTCCGCCACAG 684
QY 121 LeuLeuAspLeuAlaLeuValHisLeu-AlaValGluValAlaValThrLeuValAlaA 140
Db 683 CTGTGGATCTCGCTGCTGCTCCACCTTCGCGTGGAAGTCGCGGTTCACCTTGGTTGGGC 624
QY 140 aspaspGlnGlyHisValHisValLeuGlyPheValLeuGluAlaArgLeuGlyVa 160
Db 623 AGATGCCAGGGGCGACGTCACGTTCTCGTTGTTGGTTCTTTCGAGGCAAGACTTGGTGT 564
QY 160 lGluAspLeuLeuProGluProLeuHisLeuLeuGluGlyValAlaValValGlnThrGl 180
Db 563 CGAGGATCTGCTGCTGAGCGCTGCACCTCTCGAAGAGTTCGCGGTGTCACAGAGTGA 504
QY 180 uHisGlnAspGluAsnValSerCysGluAspGlyGluAlaProHisGlyGlyGluArgVa 200
Db 503 ACACCAAGGATGAAAACCTCTCTCTGAGGATGGAGAGCGCGCATGCGGGGAAAGGCT 444
QY 200 lValAlaGlyArgValGluAspValGluLeuValAspLeuAlaAlaAspGlyValGluLe 220
Db 443 GGTTCGCGGACGTTCTGAGGATGTCGAGCTGTAGACCTCGCGCGGATGGAGTAGAACT 384
QY 220 uAlaValGluValLeuAspGlyArgValGlyValLeuGluAlaAlaGlyGluGluAl 240
Db 383 TGGCTGGAAGTCTCTGATGTTAGGCGTGTAGCGTCTCGAAGCGCGCGGTGAGGAAGC 324
QY 240 aArgHisAspGlyArgLeuAlaHisLeuGlyArgAlaGluAspAspHisAlaAlaVa 260
Db 323 GCGACACGATGGCGCGTCTTCCACCTTGGACAGCGGAGGATGGCCATGCGATAGCACT 264
QY 260 lLeuGlyArgAspThrGlnLeuArgValAlaAlaHisLeuLeuAspHisArgGlyGl 280
Db 263 TCTTGGCGGGGATCTCAGCTCCAGTCCGCTCGGCGACATCTTCTTGATCATCGCGGCCA 204
QY 280 nPheHis 282
Db 203 GTTTCAT 197
RESULT 13
BM05574/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source

```

BM05574 1137 bp mRNA linear EST 12-MAR-2002
 AGENCOURT_6708111 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5753182
 5', mRNA sequence.
 BM05574 1 GI:19372035
 EST.
 human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1137)
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM12788 row: 1 column: 23
 High quality sequence stop: 542.
 Location/Qualifiers
 1..1137

Tue Dec 31 15:27:53 2002

Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
Washington University Genome Sequencing Center For information on
obtaining a clone please contact: Dr. Hiroshi Inoue
(hinoue@im.wustl.edu)

Seq primer: -40RP from Gibco
High quality sequence stop: 449.

FEATURES

source
1. 506
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HR85 islet"
/tissue_type="Purified pancreatic islet"
/lab_host="DH10B"
/note="Organ: Pancreas; Vector: pBluescript SK(-); Site: 1:
Not; Site 2: XhoI; CDNA made by oligo-dT priming.
Size: selected on agarose gel. Average insert size -1kb. 5'
XhoI site was destroyed after directional cloning.
Amplified once. Contact information: Hiroshi Inoue, MD,
Metabolism Div. (Alan Permutt Lab), Washington University
School of Medicine, Box 8127, 660 South Euclid Ave., St.
Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
314-362-1916, Fax: 314-747-2692."

BASE COUNT 109 a 168 c 165 g 64 t
ORIGIN

Alignment Scores: 1.66e-70 Length: 506
Pred. No.: 832.00 Matches: 168
Score: 832.00 Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 58.97% Gaps: 0
DB: 13

US-09-709-103-1F4 (1-282) x BM311047 (1-506)

QY 2 AlaAspAlaAlaLeuLeuValLeuGlyLeuAlaAlaGlyLeuLeuAlaAsp 21
Db 506 GCTGATGACGCGCTCTTGTCTTGGCTTGGCTTGGCTTGGCTTCTCGCGGAT 447
QY 22 ValHisGluValAlaValTyrAlaGlyProAlaArgGluGlyCysHisAspAlaGly 41
Db 446 GTACATGAGGTGCTGTGTGTACGTGGCGCGCGCGCGAGGGTGCCACGATGCCAAGGC 387
QY 42 ValAlaArgValAlaAlaAlaAlaAlaAlaGlyProGlnGlnLeuValProGln 61
Db 386 GTCCGCGCGGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 327
QY 62 ArgLeuLeuValGlnHisValAlaValLeuHisArgAspLeuAlaValGlnValTyrAla 81
Db 326 CGCCTTCTTGTGACACGCTGCGAGTACTGACCGAGACCTTGGCGGTGACAGTCTGGGCT 267
QY 82 HisLeuAlaGlyGlnLeuGlyHisGlyGluGluArgAlaGluHisLeuValGlnAlaAla 101
Db 266 CATCTCGTGGGACCTTGGCCATGGCGAGAGCGCGGGAACATCTGGTCCAGGCTGCT 207
QY 102 ValLeuLeuGlyArgAspLeuGluValGlyAlaAlaLeuGlyValAlaHisGlnLeu 121
Db 206 GTTCTTCTTGGCGGAGATCTCGAAGTAGCGCGCGCGCGCGCGCGCGCGCGCG 147
QY 122 LeuAspLeuAlaLeuValHisLeuAlaValGluValAlaValThrLeuValAlaAlaAsp 141
Db 146 CTGATCTCGCGTGTGTCTCCCTTCGCGGTAGAGTCCGCGGTACCCCTTGTTCGCGCAGAT 87
QY 142 AspGlnGlyHisValHisValLeuLeuGlyPheValLeuGluAlaArgLeuGlyValGlu 161
Db 86 GACGAGGCGCAGCTCCACGCTTCTCTTGGTTTGTCTTGGAGCAAGACTTGGTGTCCGAG 27
QY 162 AspLeuLeuProGluProLeuHis 169
Db 26 GATCTGCTGCTGAGCGCTGCGAC 3

Search completed: December 30, 2002, 18:04:18

QY 79 ValTrpAlaHisLeuAlaGlyGlnLeuGlyHisGlyGluGluArgAlaGluHisLeuVal 98
Db 830 GTCTGGGCTCATCTCGCTGGCAG-CTTGGCTTGGCGAGAGCGCGCGCATCTTGTGTC 772
QY 99 GlnAlaAlaValLeuLeuGlyArgAspLeuGluValGlyAlaAlaLeuGlyValValAl 118
Db 771 CAGGCTGCTGTTCTTCTTGGCGGAGATCTCCGAGTAGGCGCAGCGT---GGGTCGTGCG 715
QY 118 aHisGlnLeuLeuAsp--LeuAlaLeuValHis-LeuAlaValGluValAlaValThrLe 137
Db 714 CCACCACTGCTGATCTTCGCGCTGCTCCACCTCGCGGTAGAGTGGCGGTACACCT 655
QY 137 uValAlaAlaAspGlnGlyHisValHisValLeuLeuGlyPheValLeuGluAlaAla 157
Db 654 TGTTCGCGCAGATGACGAGGCGACGCTCCACGTTCTCTTGGTTTGTTCCTTGGAGCAAG 595
QY 157 gLeuGlyValGluAspLeuLeuProGluProLeuHisLeuLeuGluGlyValAlaValVa 177
Db 594 ACTTGGTGTGAGGATCTGCTGCTGAGCGCTGCACCTCTCGAAGGAGTCCGCGTTGT 535
QY 177 lGlnThrGlnHisGlnAspGluAsnValSerCysGluAspGlyGluAlaProHisGlyGl 197
Db 534 CGAGCTGAACACGAGATGAACAGCTCTCTGTGAGGATGGAGGCGCGCATGGCGG 475
QY 197 yGluArgValValAlaGlyArgValGluAspValGluLeuValAspLeuAlaAlaAspGl 217
Db 474 GGAACGGGTGTTCGCGACGTGTGAGGATGTTCAGCTGTGTAGACCTCGCGCGGATGG 415
QY 217 yValGluLeuAlaValGluValLeuAspGlyArgValGlyValGluValGluAlaGl 237
Db 414 AGTAGACTTGGGTGGAAGTCTCGATGTGTAGCGGTAGGCGTCTCGAAGCGCGCGG 355
QY 237 yGluGluAlaArgHisAspGlyArgLeuAlaHisLeuGlyArgAlaGluAspHisAl 257
Db 354 TGAGGAAGCGCGACACGATGGCGCTTTCGCCACCTTGGACGCGCGAGATGACCATGC 295
QY 257 aLeuAlaValLeuGlyArgAspThrGlnLeuArgValAlaAlaAlaHisLeuLeuAspHi 277
Db 294 GATAGCAGTTCTTGGCGGGGATCTCAGCTCGGATGCTCGGCGACATCTTCTTGTATCA 235
QY 277 sArgGly-GlnPheHis 282
Db 234 TCGCGCGCCAGTTTCAT 218

RESULT 15
BM311047/c 506 bp mRNA linear EST 03-JAN-2002
LOCUS ig60a06.y1 HR85 islet Homo sapiens cDNA 5' similar to TR:Q9Y272
DEFINITION Q9Y272 RAS-RELATED PROTEIN. i, mRNA sequence.

ACCESSION BM311047
VERSION BM311047.1 GI:18045392
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 506)
AUTHORS Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K.,
Lemishka, I., Seearce, M., Brestelli, J., Gradwohl, G., Clifton, S.,
Hallier, L., Matra, M., Pape, D., Wylie, T., Martin, J., Blistain, A.,
Schmitt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J., Cardenas
M., Gibbons, M., McCann, R., Cole, R., Tsagareishvili, R., Williams, T.,
Jackson, Y. and Bowers, Y.

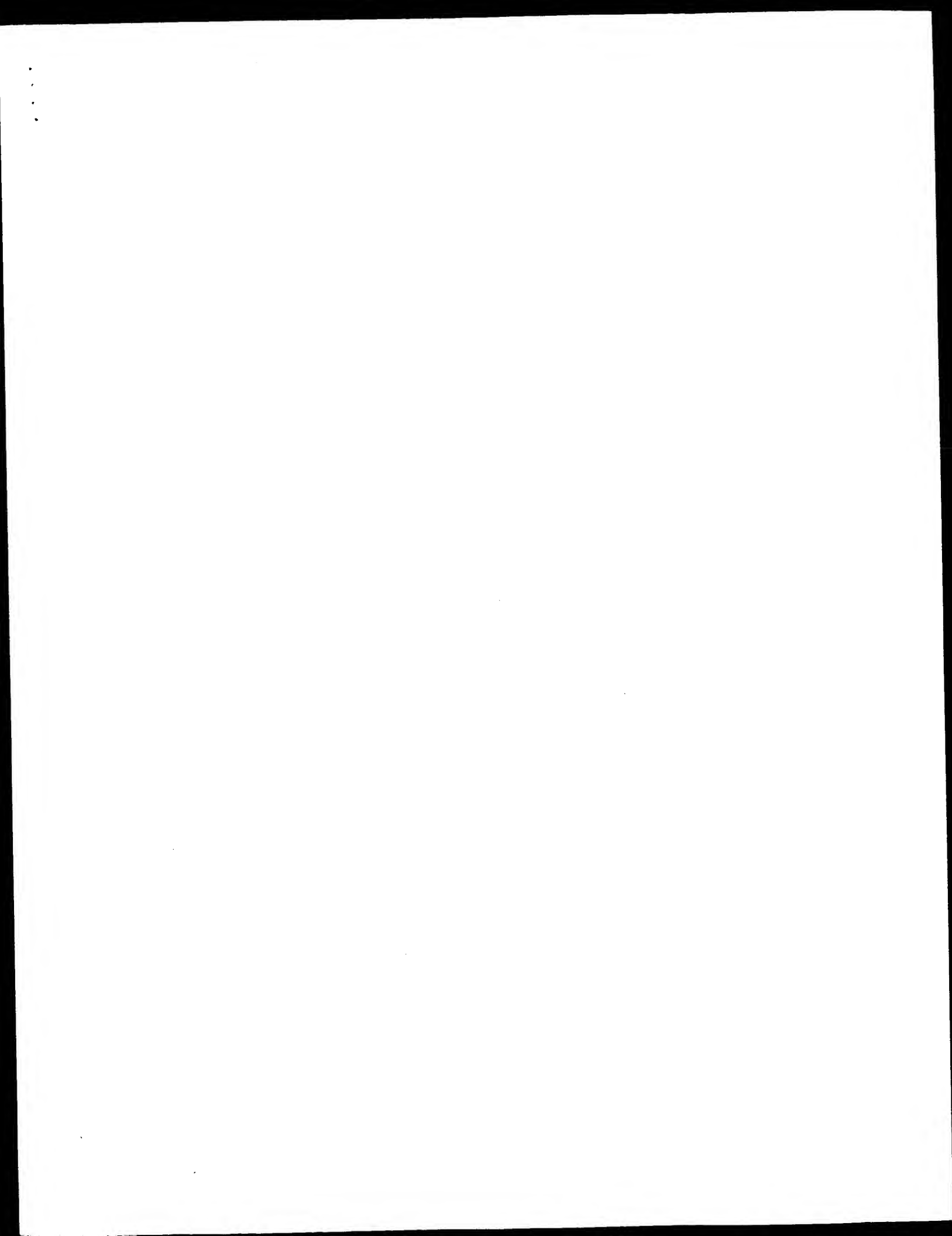
TITLE Endocrine Pancreas Consortium
JOURNAL Unpublished (2000)
COMMENT Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138

Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@bioh.p.harvard.edu

Tue Dec 31 15:27:53 2002

us-09-709-103-1f4.p2n.rst

Job time : 2185 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: December 30, 2002, 14:46:00 ; Search time 355 Seconds

(without alignments)
1788.911 Million cell updates/sec

Title: US-09-709-103-1f4

Sequence: 1 LADDAALLVGLAAGAGLLA.....RTQLRVARAHLDHKGQFH 282

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 segs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n model -DEV=xlp
-O/cgnt_1/USFto.spool/SULLVNA1f4/runat_30122002_144046_21490/app.query.fasta_1.455
-DB=NGeneSeq_101002 -QFMT=fastcap -SUFFIX=p2n.rng -MINMATCH=0.1 -IOPCT=0
-LOOEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USBR=SULLVNA1f4 @CGN 1.1 79 @runat_30122002_144046_21490 -NCPU=6 -ICPU=3
-NO XLPY -NO MMF -IARSEQOVRT -NEG_SCORES=0 -WAIT -DONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-XGAPOP=10 -XGAPEXT=0.5 -DELop=6 -DELEXT=7

Database :

N GeneSeq_101002: *
1: /SID2/gcgdata/geneSeq/geneSeqn-emb1/NA1980.DAT: *
2: /SID2/gcgdata/geneSeq/geneSeqn-emb1/NA1981.DAT: *
3: /SID2/gcgdata/geneSeq/geneSeqn-emb1/NA1982.DAT: *
4: /SID2/gcgdata/geneSeq/geneSeqn-emb1/NA1983.DAT: *
5: /SID2/gcgdata/geneSeq/geneSeqn-emb1/NA1984.DAT: *
6: /SID2/gcgdata/geneSeq/geneSeqn-emb1/NA1985.DAT: *
7: /SID2/gcgdata/geneSeq/geneSeqn-emb1/NA1986.DAT: *
8: /SID2/gcgdata/geneSeq/geneSeqn-emb1/NA1987.DAT: *
9: /SID2/gcgdata/geneSeq/geneSeqn-emb1/NA1988.DAT: *
10: /SID2/gcgdata/geneSeq/geneSeqn-emb1/NA1989.DAT: *
11: /SID2/gcgdata/geneSeq/geneSeqn-emb1/NA1990.DAT: *
12: /SID2/gcgdata/geneSeq/geneSeqn-emb1/NA1991.DAT: *
13: /SID2/gcgdata/geneSeq/geneSeqn-emb1/NA1992.DAT: *
14: /SID2/gcgdata/geneSeq/geneSeqn-emb1/NA1993.DAT: *
15: /SID2/gcgdata/geneSeq/geneSeqn-emb1/NA1994.DAT: *
16: /SID2/gcgdata/geneSeq/geneSeqn-emb1/NA1995.DAT: *
17: /SID2/gcgdata/geneSeq/geneSeqn-emb1/NA1996.DAT: *
18: /SID2/gcgdata/geneSeq/geneSeqn-emb1/NA1997.DAT: *
19: /SID2/gcgdata/geneSeq/geneSeqn-emb1/NA1998.DAT: *
20: /SID2/gcgdata/geneSeq/geneSeqn-emb1/NA1999.DAT: *
21: /SID2/gcgdata/geneSeq/geneSeqn-emb1/NA2000.DAT: *
22: /SID2/gcgdata/geneSeq/geneSeqn-emb1/NA2001.DAT: *
23: /SID2/gcgdata/geneSeq/geneSeqn-emb1/NA2001B.DAT: *
24: /SID2/gcgdata/geneSeq/geneSeqn-emb1/NA2002.DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
C 1	1411	100.0	846	21	AAZ36892
C 2	1411	100.0	1740	21	AAZ36914
C 3	1411	100.0	1776	21	AAZ36917
C 4	1411	100.0	1801	21	AAZ36893
C 5	1411	100.0	1841	20	AAZ36922
C 6	1348	95.5	3986	20	AAZ36923
C 7	963.5	68.3	1689	20	AAZ36924
C 8	911	64.6	3079	20	AAZ36925
C 9	579.5	41.1	837	21	AAZ36913
C 10	579.5	41.1	2832	24	ABL92076
C 11	579.5	41.1	2973	24	ABL92087
C 12	579.5	41.1	3427	24	ABK71553
C 13	498.5	35.3	3020	24	ABJ92134
C 14	478.5	33.9	951	23	AAJ90571
C 15	422	29.9	624	24	ABQ28318
C 16	422	29.9	624	24	ABQ28319
C 17	321	22.7	624	24	ABQ28320
C 18	321	22.7	624	24	ABQ28321
C 19	267	18.9	1305	23	ABJ07789
C 20	244.5	17.3	3061	24	ABQ72648
C 21	228.5	16.2	1249	24	ABQ72525
C 22	219	15.5	368	22	ABA51382
C 23	219	15.5	368	22	ABA69388
C 24	219	15.5	368	22	ABA69383
C 25	219	15.5	368	22	AAK17661
C 26	219	15.5	368	22	AAK43477
C 27	219	15.5	368	22	AAI24259
C 28	219	15.5	368	22	AAI49541
C 29	219	15.5	368	22	AAI49541
C 30	219	15.5	368	22	ABJ16008
C 31	188.5	13.4	1144	22	AAH90075
C 32	188.5	13.4	4933	22	AAK81235
C 33	188.5	13.4	4933	22	AAK81234
C 34	184.5	13.1	1087	22	AAI60838
C 35	184.5	13.1	1108	22	AAI59052
C 36	184.5	13.1	4167	24	AAD37605
C 37	166.5	11.8	702	23	ABL14159
C 38	166	11.8	3474	14	AAQ39018
C 39	166	11.8	3474	18	AAT74012
C 40	166	11.8	3474	19	AAV16191
C 41	161	11.4	6338	24	ABK91481
C 42	161	11.4	6610	24	ABK91514
C 43	161	11.4	789	24	ABK91496
C 44	159.5	11.3	2322	21	AAA51961
C 45	159.5	11.3	2565	24	ABL40018

ALIGNMENTS

RESULT 1
AAZ36892/c
AAZ36892 standard; cDNA: 846 BP.
AAZ36892;
13-MAR-2000 (first entry)
cDNA encoding an activator of G protein signalling (AGS) protein.
Activator of G protein signalling; AGS; ras-related G protein;
GTP hydrolysis; G protein activity; pheromone response pathway;
G protein-coupled signal transduction; G-gamma selectivity;
cellular signal transduction; ss.
Homo sapiens.
Location/Qualifiers
Key

Accession	Species	Protein Name	Length	Matches	Conservative	Mismatches	Indels	Gaps
Q09958	Human	Protein kinase C- δ	846	846	282	0	0	0
Q09959	Human	Protein kinase C- δ	846	846	282	0	0	0
Q09960	Human	Protein kinase C- δ	846	846	282	0	0	0
Q09961	Human	Protein kinase C- δ	846	846	282	0	0	0
Q09962	Human	Protein kinase C- δ	846	846	282	0	0	0
Q09963	Human	Protein kinase C- δ	846	846	282	0	0	0
Q09964	Human	Protein kinase C- δ	846	846	282	0	0	0
Q09965	Human	Protein kinase C- δ	846	846	282	0	0	0
Q09966	Human	Protein kinase C- δ	846	846	282	0	0	0
Q09967	Human	Protein kinase C- δ	846	846	282	0	0	0
Q09968	Human	Protein kinase C- δ	846	846	282	0	0	0
Q09969	Human	Protein kinase C- δ	846	846	282	0	0	0
Q09970	Human	Protein kinase C- δ	846	846	282	0	0	0
Q09971	Human	Protein kinase C- δ	846	846	282	0	0	0
Q09972	Human	Protein kinase C- δ	846	846	282	0	0	0
Q09973	Human	Protein kinase C- δ	846	846	282	0	0	0
Q09974	Human	Protein kinase C- δ	846	846	282	0	0	0
Q09975	Human	Protein kinase C- δ	846	846	282	0	0	0
Q09976	Human	Protein kinase C- δ	846	846	282	0	0	0
Q09977	Human	Protein kinase C- δ	846	846	282	0	0	0
Q09978	Human	Protein kinase C- δ	846	846	282	0	0	0
Q09979	Human	Protein kinase C- δ	846	846	282	0	0	0
Q09980	Human	Protein kinase C- δ	846	846	282	0	0	0
Q09981	Human	Protein kinase C- δ	846	846	282	0	0	0
Q09982	Human	Protein kinase C- δ	846	846	282	0	0	0
Q09983	Human	Protein kinase C- δ	846	846	282	0	0	0
Q09984	Human	Protein kinase C- δ	846	846	282	0	0	0
Q09985	Human	Protein kinase C- δ	846	846	282	0	0	0
Q09986	Human	Protein kinase C- δ	846	846	282	0	0	0
Q09987	Human	Protein kinase C- δ	846	846	282	0	0	0
Q09988	Human	Protein kinase C- δ	846	846	282	0	0	0
Q09989	Human	Protein kinase C- δ	846	846	282	0	0	0
Q09990	Human	Protein kinase C- δ	846	846	282	0	0	0
Q09991	Human	Protein kinase C- δ	846	846	282	0	0	0
Q09992	Human	Protein kinase C- δ	846	846	282	0	0	0
Q09993	Human	Protein kinase C- δ	846	846	282	0	0	0
Q09994	Human	Protein kinase C- δ	846	846	282	0	0	0
Q09995	Human	Protein kinase C- δ	846	846	282	0	0	0
Q09996	Human	Protein kinase C- δ	846	846	282	0	0	0
Q09997	Human	Protein kinase C- δ	846	846	282	0	0	0
Q09998	Human	Protein kinase C- δ	846	846	282	0	0	0
Q09999	Human	Protein kinase C- δ	846	846	282	0	0	0
Q10000	Human	Protein kinase C- δ	846	846	282	0	0	0
Q10001	Human	Protein kinase C- δ	846	846	282	0	0	0

101	Qy	AlaValLeuLeuGlyArgAspLeuGluValGlyAlaAlaLeuGlyValValAlaHisGln	120
546	Ddb	CTGTGTTCTTCTGGCCGAGATCTCGAAGTAGGCGCAGCGCTGGGGGTCTGTGCCCCACCAG	487
121	Qy	LeuLeuAspLeuAlaLeuValHisLeuAlaValGluValAlaValThrLeuValAlaAla	140
486	Ddb	CTGCTCGATCTTCGCGTGCTGCACCTCGCGGTAGAAGTCGCGGTACACCTTTGTTCCCGCA	427
141	Qy	AspAspGlnGlyHisValHisValLeuLeuGlyPheValLeuGluAlaArgLeuGlyVal	160
426	Ddb	GATGACCAGGGGGCAGCTCCACGTTCTCTTGGTTTGTCTTGAGGCCAAGACTTGGTGTCTC	367
161	Qy	GluAspLeuLeuProGluProLeuHisLeuLeuGlyValAlaValValGlnThrGlu	180
366	Ddb	GAGGATCTGTGCTGCTGAGCGCTGCACCTCTCGAAGGAGTCGCGTGTTCGAGACTGAA	307
181	Qy	HisGlnAspGluAsnValSerCysGluAspGlyGluAlaProHisGlyGlyGluArgVal	200
306	Ddb	CACCAGAGTAGAAACAGTCCTCTGTGAGATGTGAGAGCGCCGCATGCGGGGAAACGGGTG	247
201	Qy	ValAlaGlyArgValGluAspValGluLeuValAspLeuAlaAlaAspGlyValGluLeu	220
246	Ddb	GTTCGCCGACGTTGTCGAGGATGTTCGAGCTGTGTAGACCTCGCCGCGAGTAGAAGT	187
221	Qy	AlaValGluValLeuAspGlyArgValGlyValLeuGluAlaAlaGlyGlyGluGluAla	240
186	Ddb	GGCGTGGAAAGTCCTCGATGTGTAGCGCTCTCGAAGCGCGCGGTGAGGAAGCG	127
241	Qy	ArgHisAspGlyArgLeuAlaHisLeuGlyArgAlaGluAspAspHisAlaIleAlaVal	260
126	Ddb	CGACACGATGCGCGCTCTTGCCCACTTTGGACGAGCCGAGGATGACCATGCGATAGCAGTT	67
261	Qy	LeuGlyArgAspThrGlnLeuArgValAlaAlaArgAlaHisLeuLeuAspHisArgGlyGln	280
66	Ddb	CTTGGCCGGGATCTCACTCCGAGTCCGTCGCGGCACATCTTCTTGATCATCGCGGCCAG	7
281	Qy	PheHis	282
6	Ddb	TTCAT	1
RESULT 2			
ID	AAZ36914/c		
ID	AAZ36914 standard; DNA; 1740 BP.		
XX	AAZ36914;		
AC	AAZ36914;		
XX	13-MAR-2000 (first entry)		
DT	DNA encoding an activator of G protein signalling (AGS) protein.		
DE	Activator of G protein signalling; AGS; ras-related G protein;		
XX	GTP hydrolysis; G protein activity; pheromone response pathway;		
KW	G protein-coupled signal transduction; G-gamma selectivity;		
KW	cellular signal transduction; ss.		
XX	Homo sapiens.		
OS	Location/Qualifiers		
XX	145..991		
FT	/*tag= a		
FT	/product= "activator of G protein signalling (AGS) protein"		
FT			
XX	WO9958670-A1.		
PN	18-NOV-1999.		
PD	07-MAY-1999;		
XX	99WO-US10151.		
PF			
XX	08-MAY-1998;		
XX	98US-0084842.		
PR	07-OCT-1998;		
PR	98US-0103355.		
XX	(CADU-) CADUS PHARM CORP.		

CC expression.

SQ Sequence 1801 BP; 437 A; 561 C; 500 G; 303 T; 0 other;

Alignment Scores:

Pred. No.:	1,06e-88	Length:	1801
Score:	1411.00	Matches:	282
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	21	Gaps:	0

US-09-709-103-1f4 (1-282) x AA236893 (1-1801)

```

QY 1 LeuAlaAspAspAlaAlaLeuLeuValLeuGlyLeuAlaAlaGlyAlaGlyLeuLeuAla 20
Db 999 CTAGCTGATGACGACGCGCTCTTGTCTTGGCTGGCTGGCGGCTTCTTCTGCG 940
QY 21 AspValHisGluValAlaValTyrAlaGlyProAlaArgGluGlyCysHisAspAlaLys 40
Db 939 GATGTACATGAGAGTCTGCTGTACCTGGCGCGCGCGCGCGAGAGGTGCGACGATGCCAA 880
QY 41 GlyValAlaArgValAlaAlaAlaAlaAlaAlaAlaGlyProGlnGlnLeuLeuValPro 60
Db 879 GCGCGTCCCGCGGCTGCGCGCGCGCGCGCGCTGCGCGCGCGCGCGCTTGTCTTCCG 820
QY 61 GlnArgLeuLeuValGlnHisValAlaValLeuHisArgAspLeuAlaValGlnValTyr 80
Db 819 CAGCGCCTTCTTGTGACACGACGCTGCACTGCAACGAGCCTTGGCGGTGAGGTCTCG 760
QY 81 AlaHisLeuAlaGlyGlnLeuGlyHisGlyGlyGluTyrAlaGlnHisLeuValGlnAla 100
Db 759 GCTGATCTCGCTGGGAGCTTGGCCATGGCGAGAGCGCGGAGAACCTTGTCTGCGAGCT 700
QY 101 AlaValLeuLeuGlyArgAspLeuGluValGlyAlaAlaLeuGlyValAlaHisGln 120
Db 699 CCTGTTCTTCTTGGCCAGATCTCGAAGTAGGCGCGCGCGCGGCTGCGCGCGCGCGCG 640
QY 121 LeuLeuAspLeuAlaLeuValHisLeuAlaValGluValAlaValHisLeuValAlaAla 140
Db 639 CTGCTTCATCTCGCGCTGCTGCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 580
QY 141 AspArgGlnGlnHisValHisValLeuLeuGlyPheValLeuGlnAlaArgLeuGlyVal 160
Db 579 GATGACCAAGGGGACGCTCCAGTCTCTTGGTTTGTCTTGGGCAAGACTTGTGTCTC 520
QY 161 GlnAspLeuLeuProGluProLeuHisLeuLeuGlnGlyValAlaValAlaGlnThrGln 180
Db 519 GAGGATCTGCTGCTGACGCGCTGCACTCTCTGGAAGAGTGGCGGTTGTCCAGACTGAA 460
QY 181 HisGlnAspGlnAsnValSerCysGluAspGlyGluAlaProHisGlyGlyGluArgVal 200
Db 459 CACCAGAGATGAAAACGCTCTCTGTGAGATGAGAGCGCGCGCGCGAGGAGGAGGAGG 400
QY 201 ValAlaGlyArgValGluGlnValGluLeuValAspLeuAlaAlaAspGlyValGluLeu 220
Db 399 GTTGGCGGACGCTGCTGAGGATGTCAAGCTGTGAACCTGCGCGCGAGTGGAGAACTT 340
QY 221 AlaValGluValLeuAspGlyArgArgValGlyValLeuGlnAlaAlaGlyGlnGluAla 240
Db 339 GCGGTGGAAGTCTCGATGTGAGGTGAGCGCTCTCGAAGGCGCGCGAGGAAAGCG 280
QY 241 ArgHisAspGlyArgLeuAlaHisLeuGlyArgAlaGluAspAspHisAlaAlaValVal 260
Db 279 CACACGATGCGCGCTTGTGCGCACCTTGGACGAGCGCGAGTGCACATGGGATGCGATT 220
QY 261 LeuGlyArgAspThrGlnLeuArgValAlaArgAlaHisLeuLeuAspHisArgGlyGln 280
Db 219 CTGGCGCGGATCTCACTCGAAGTGGCTGGGCAATCTTGTGATCATCGCGGCGAG 160
QY 281 PheHis 282
Db 159 TTTCAAT 154

```

RESULT 5

AA23022/C

ID AA23022 standard; cDNA, 1841 BP.

XX AC AA23022;

XX AC 17-JAN-2000 (first entry)

XX DE Human kd312 polypeptide encoding cDNA.

XX KW kd312 polypeptide; cancer; immunodeficiency disorder; AIDS; stroke;

XX KW heart attack; head trauma; neurodegenerative disease; human;

XX KW Parkinson's disease; Alzheimer's disease; ss.

XX OS Homo sapiens.

XX PN WO950288-A2.

XX PD 07-OCT-1999.

XX PF 30-MAR-1999; 99MO-US06993.

XX PR 31-MAR-1998; 98US-0053374.

XX PA (AMGE-) AMGEN INC.

XX PI Yen K;

XX XX WPI; 1999-601322/51.

XX DR P-PSDB; AAY42693.

XX PT kd312 polypeptides useful for treating diseases and disorders

XX PT associated with alterations in cell proliferation and cell death

XX PS Claim 1; Fig 8; 85p; English.

XX CC The invention provides nucleic acid molecules encoding human and rat

XX CC kd312 polypeptides. The kd312 polypeptides can be expressed by standard

XX CC recombinant methodology. The kd312 sequences, and the antibodies against

XX CC the proteins may be used to treat or diagnose the presence or progression

XX CC of various diseases, e.g. cancer, immunodeficiency disorders (e.g. AIDS),

XX CC stroke, heart attack, head trauma, and neurodegenerative diseases (e.g.

XX CC Parkinson's disease and Alzheimer's disease). The present sequence

XX CC represents the human kd312 cDNA sequence.

SQ Sequence 1841 BP; 398 A; 605 C; 530 G; 308 T; 0 other;

Alignment Scores:

Pred. No.:	1.09e-88	Length:	1841
Score:	1411.00	Matches:	282
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	20	Gaps:	0

US-09-709-103-1f4 (1-282) x AA23022 (1-1841)

```

QY 1 LeuAlaAspAspAlaAlaLeuLeuValLeuGlyLeuAlaAlaGlyAlaGlyLeuLeuAla 20
Db 1100 CTAGCTGATGACGACGCGCTCTTGTCTTGGCTGGCTGGCGGCTTCTTCTGCG 1041
QY 21 AspValHisGluValAlaValTyrAlaGlyProAlaArgGluGlyCysHisAspAlaLys 40
Db 1040 GATGTACATGAGTCTGCTGTGATGCTGGGCGCGCGCGCGAGAGGTGCGACATGCCAA 981
QY 41 GlyValAlaArgValAlaAlaAlaAlaAlaAlaAlaGlyProGlnGlnLeuLeuValPro 60
Db 980 GCGGTGCGCGGCTGCGCGCGCGCGCGCGCTGCGCGCGCGCGCGAGCTTGTCTTCCG 921
QY 61 GlnArgLeuLeuValGlnHisValAlaValLeuHisArgAspLeuAlaValGlnValTyr 80
Db 920 CAGCGCCTTCTTGTGACGACGCTGCACTGCAACGAGCCTTGGCGGTGCGGTCTGG 861

```

us-09-709-103-1f4.p2n.rng

Tue Dec 31 15:27:52 2002

```

DR WPI: 1999-601322/51.
DR P-PSDB; AAY42693.
XX kd312 polypeptides useful for treating diseases and disorders
PT associated with alterations in cell proliferation and cell death
XX Claim 1; Fig 5; 85pp; English.
XX The invention provides nucleic acid molecules encoding human and rat
CC kd312 polypeptides. The kd312 polypeptides can be expressed by standard
CC recombinant methodology. The kd312 sequences, and the antibodies against
CC the proteins may be used to treat or diagnose the presence or progression
CC of various diseases, e.g. cancer, immunodeficiency disorders (e.g. AIDS),
CC stroke, heart attack, head trauma, and neurodegenerative diseases (e.g.
CC Parkinson's disease and Alzheimer's disease). The present sequence
CC represents the human kd312 genomic DNA sequence.
XX Sequence 3986 BP; 856 A; 1236 C; 1172 G; 732 T; 0 other;

Alignment Scores:      5,19e-84      Length:      3986
Pred. No.:      1348.00      Matches:      282
Score:      79.89%      Conservative:      0
Percent Similarity:      79.89%      Mismatches:      0
Best Local Similarity:      95.54%      Indels:      71
Query Match:      20      Gaps:      1
DB:

US-09-709-103-1F4 (1-282) x AA223023 (1-3986)
QY 1 LeuAlaAspAspAlaAlaLeuLeuValLeuGlyLeuAlaAlaGlyAlaGlyLeuAla 20
Db 1830 CTAGCTGATGACGACGCTCTTGTCTTGGCTGCTGCTGCTGCTGCTGCTGCTG 1771
QY 21 AspValHisGluValAlaValTyAlaGlyProAlaArgGluGlyCysHisAspAlaLys 40
Db 1770 GATGTACATGAGTGCCTGTGTACGTGGCGCGCGCGCGCGCGCGCGCGCGCG 1711
QY 41 GlyValAlaAlaArgValAlaAlaAlaAlaAlaAlaAlaGlyProGlnGlnLeuValPro 60
Db 1710 GCGTTCGCGCGCGCGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1651
QY 61 GlnArgLeuValGlnHisValAlaValLeuHisArgAspLeuAlaValGlnValTrp 80
Db 1650 CAGCGCTCTTCTTGTGCGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1591
QY 81 AlaHisLeuAlaGlyGlnLeuGlyHisGlyGluAlaArgAlaGluHisLeuValGlnAla 100
Db 1590 GCTCATCTGCTGGCGAGCTTGGCCATGCGAAGAGCGCGCGAACAATCTGCTCAG 1531
QY 101 AlaValLeuGlyArgAspLeuGluValGlyAlaAlaLeuGlyValValAlaHisGln 120
Db 1530 GCTGTTCTTCTTGGCGAGATCTCGAAGTAGGCGCGCGCTGGGGTCTGCTGCC 1471
QY 121 LeuLeuAspLeuAlaLeuValHisLeuAlaValGluValAlaValThrLeuValAlaAla 140
Db 1470 CTGCTCATCTGCGCTGGTTCACCTTCGCGGTAGAGTCCGCTCACCCTTCTTGC 1411
QY 141 AspAspGlnGlyHisValHisValLeuLeuGlyPheValLeuGluAlaArgLeuGlyVal 160
Db 1410 GATGCCAGGGGCGACGCTCCACGTTCTTGTGTTTGTCTTGGAGCAAGACTTGT 1351
QY 161 GluAspLeuLeuProGluProLeuHisLeuGluGlyValAlaValValGlnThrGlu 180
Db 1350 GAGGATCTGCTGCTGAGCGCGCTGCACCTCTCTGAGGAGTCCGCGTGTGTCAG 1291
QY 181 HisGlnAspGluAsnValSer----- 187
Db 1290 CACCAAGGATGAACGTTCTCTAGAGGGGCGACAGAGAGAGAGAGAGAGAGAG 1231
QY 187 ----- 187
Db 1230 TGGCGCGCGCGAGGGGCAAGTCTGAGGCTGACTTTGAGCGCGCGCGCGCTAG 1171

```

```

QY 81 AlaHisLeuAlaGlyGlnLeuGlyHisGlyGluGluArgAlaGluHisLeuValGlnAla 100
Db 860 GCTCATCTGCTGCGGAGCTTGGCCATGGCGAAGAGCGCGCGGACATCTGCTCAGGCT 801
QY 101 AlaValLeuLeuGlyArgAspLeuGluValGlyAlaAlaLeuGlyValValAlaHisGln 120
Db 800 GCTGTTCTTCTTGGCCGAGATCTCGAAGTAGGCGCGAGCTGCGGGTCTGCTCC 741
QY 121 LeuLeuAspLeuAlaLeuValHisLeuAlaValAlaValAlaValThrLeuValAlaAla 140
Db 740 CTGCTCATCTGCGCTGCTGCTCCACCTCGCGGTAGAGTCCGCTCACCCTTGTTC 681
QY 141 AspAspGlnGlyHisValHisValLeuLeuGlyPheValLeuGluAlaArgLeuGlyVal 160
Db 680 GATGCCAGGGGCGACGCTCCACGTTCTTGTGTTTGTCTTGGAGCAAGACTTGT 621
QY 161 GluAspLeuLeuProGluProLeuHisLeuLeuGluGlyValAlaValValGlnThrGlu 180
Db 620 GAGGATCTGCTGCTGAGCGCTGCACCTCTCGAAGGAGTGCAGGTTGTCCAGACT 561
QY 181 HisGlnAspGluAsnValSerCysGluAspGlyGluAlaProHisGlyGlyGluArgVal 200
Db 560 CACCAAGGATGAACGTTCTTGTGAGGATGAGAGCGCGCGCTGCGGGGAAACGG 501
QY 201 ValAlaGlyArgValGluAspValGluLeuValAspLeuAlaAlaAspGlyValGluLeu 220
Db 500 GTTCCGCGAGCTGCTGAGGATCTGAGCTGTGAGTCTGCGCGGATGAGAGTGA 441
QY 221 AlaValGluValLeuAspGlyArgValGlyValLeuGluAlaAlaGlyGluGluAla 240
Db 440 GCGGTGGAAGTCTCATGATGAGGCTGTAGCGCTGCTCGAAGCGCGCGGTGAG 381
QY 241 ArgHisAspGlyArgLeuAlaHisLeuGlyArgAlaGluAspAspHisAlaAlaVal 260
Db 380 CGACACATGCGCGCTTGTGCCACCTTGGACGAGCGGAGGATGACCATGCGAT 321
QY 261 LeuGlyArgAspThrGlnLeuArgValAlaArgAlaHisLeuLeuAspHisArgGlyGln 280
Db 320 CTTGGCGGGGATATCAGCTCCGAGTCCGCTGCGGGCACATCTTCTTGATCAT 261
QY 281 PheHis 282
Db 260 TTTTAT 255

RESULT 6
AA223023/c
ID AA223023 standard; DNA; 3986 BP.
XX AC AA223023;
XX DT 17-JAN-2000 (first entry)
XX DE Human kd312 genomic DNA sequence.
XX KW kd312 polypeptide; cancer; immunodeficiency disorder; AIDS; stroke;
XX KW heart attack; head trauma; neurodegenerative disease; human;
XX KW Parkinson's disease; Alzheimer's disease; ss.
XX OS Homo sapiens.
XX PN WO950288-A2.
XX XX 07-OCT-1999.
XX XX 30-MAR-1999; 99WO-US06993.
XX XX 31-MAR-1998; 98US-0053374.
XX XX (AMGE-) AMGEN INC.
XX XX Yen K;
XX XX

```

QY 187 ----- 187
 Db 1170 TAAGCGGAAGCGCACTACGCGCGCAGACGCCGAGCCGGCTCGGGGCGCCCTG 1111
 QY 188 ----- CysGluAs 190
 Db 1110 GCGAGGGTTCCTCCGCTTCTCCCGACCTGCCCGCCCGCTCAGCTGTAGAGA 1051
 QY 190 pGlyGluAlaProHisGlyGlyGluArgValAlaIleArgValGluAspValGluLe 210
 Db 1050 TGGAGAGCGCCCGCATGCGCGGAGACGGGTGGTTCGCGACCTGTGAGATGTGAGCT 991
 QY 210 vAlaAspLeuAlaAlaAspGlyValGluLeuAlaValGluAlaAspGlyArgArgVa 230
 Db 990 GGTAGACTTCGCGCGAGATGAGTACCTTCGCGTGAAGTCTCGATGTGAGCGCTG 931
 QY 230 lGlyValLeuGluAlaAlaGlyGlyGluAlaArgHisAspGlyArgLeuAlaHisLeuG 250
 Db 930 AGGCTCTCTCGAAGCGCGCGGTGAGAGACCGGACACGATGGCTGTGGCCCACTTGG 871
 QY 250 yArgAlaGluAspAspHisAlaIleAlaValLeuGlyArgAspThrGlnLeuArgValAl 270
 Db 870 ACCAGCGCGAGATGACCAATGCGATGACAGTCTTGGCGGGATCTCAGCTCCAGATGCG 811
 QY 270 aArgAlaHisLeuLeuAspHisAspGlyGlnPheHis 282
 Db 810 TCGGGCACTCTTCTTGATCATCGCGCCAGTTTCAT 774
 RESULT 7
 AA23024/C
 ID AA23024 standard; cDNA; 1689 BP.
 AC AA23024;
 XX 17-JAN-2000 (first entry)
 DT
 DE Rat kd312 polypeptide encoding cDNA.
 XX
 KW kd312 polypeptide; cancer; immunodeficiency disorder; AIDS; stroke;
 KW heart attack; head trauma; neurodegenerative disease; rat;
 KW Parkinson's disease; Alzheimer's disease; ss.
 OS Rattus sp.
 XX
 PN W09950288-A2.
 PD 07-OCT-1999.
 XX
 PF 30-MAR-1999; 99MO-US06993.
 XX
 PR 31-MAR-1998; 98US-0053374.
 XX
 PA (AMGE-) AMGEN INC.
 XX
 PI Yen K.
 XX
 DR WPI; 1999-601322/51.
 DR P-PSDB; AAY42694.
 XX
 PT kd312 polypeptides useful for treating diseases and disorders
 associated with alterations in cell proliferation and cell death
 XX
 PS Claim 2; Fig 9; 85pp; English.
 XX
 CC The invention provides nucleic acid molecules encoding human and rat
 CC kd312 polypeptides. The kd312 polypeptides can be expressed by standard
 CC recombinant methodology. The kd312 sequences, and the antibodies against
 CC the proteins may be used to treat or diagnose the presence or progression
 CC of various diseases, e.g. cancer, immunodeficiency disorders (e.g. AIDS),
 CC stroke, heart attack, head trauma, and neurodegenerative diseases (e.g.
 CC Parkinson's disease and Alzheimer's disease). The present sequence
 CC represents the rat kd312 cDNA sequence.

SQ Sequence 1689 BP; 417 A; 462 C; 434 G; 376 T; 0 other;
 Alignment Scores:

Pred. No.:	7,1e-58	Length:	1689
Score:	963.50	Matches:	200
Percent Similarity:	79.79%	Conservative:	25
Best Local Similarity:	70.92%	Mismatches:	56
Query Match:	68.28%	Indels:	1
DB:	20	Gaps:	1

US-09-709-103-1F4 (1-282) x AA23024 (1-1689)

QY 1 LeuAlaAspAspAlaAlaLeuLeuValLeuGlyLeuAlaAlaGlyValAlaGlyLeuLeuAla 20
 Db 974 CTAACTGATGACACAGCCCTCTTCTCTTACCTGGCTGTGACACTGTTTCTTCACG 915
 QY 21 AspValHisGluValAlaValAlaValAlaValAlaValAlaValAlaValAlaValAlaVal 40
 Db 914 AATGTACATGAGGTGCTATGACGCTAGGCTCGGACAAAGGGCGCCAAAGATGCCAA 855
 QY 41 GlyValAlaArgValAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 60
 Db 854 GGCATCTCC---GTGGTGGCCCCCACCCTCCGCTCCCGCAGCAGAAAGCTTCTGTTCCT 798
 QY 61 GlnArgLeuLeuValGlnHisValAlaValLeuHisArgAspLeuAlaValGlnValAlaTyr 80
 Db 797 CAGAGCTTTTGTGACAGCAGTCACTGACAGACCACTTGGGGTGCAGATCCAG 738
 QY 81 AlaHisLeuAlaGlyGlnLeuGlyHisGlyGluGluArgAlaGluHisLeuValGlnAla 100
 Db 737 GCTCATCTCGCTAGCAGCTTGGCCATGGCAAAAGCGACAGAAATCTGTCTCAGGCT 678
 QY 101 AlaValLeuLeuGlyArgAspLeuGlyValGlyAlaAlaLeuGlyValAlaValAlaHisGln 120
 Db 677 GCTATCTCTTGGCCGAGATCTCGAAAGTAGGCACAACGCTAGAGGTCATTCGCCACAG 618
 QY 121 LeuLeuAspLeuAlaLeuValHisLeuAlaValGluValAlaValThrLeuValAlaAla 140
 Db 617 CTGCTCAATCTCCCGCTGCTCCACTTCGCGGTAGAGTCCCGCTCCCTTGTACCGCA 558
 QY 141 AspAspGlnGlyHisValHisValLeuLeuGlyPheValLeuGluAlaArgLeuGlyVal 160
 Db 557 AATGACGAGCGCAGCCACATCTCTTGGTTTGTCTTGAAGACAGACTTGTGTG 498
 QY 161 GluAspLeuLeuProGluProLeuHisLeuLeuGlyGlyValAlaValAlaValGlnThrGlu 180
 Db 497 TAGGATCTGCTGTTAGGCTTTGACCTCTCGAAGGAGTGGCGGTTGTCTAAGCTGAA 438
 QY 181 HisGlnAspGluAsnValSerCysGluAspGlyGluAlaProHisGlyGlyGluArgVal 200
 Db 437 CACCAGATGAAACCTCTCTCTGAGATAGAGAGCGCCGATGAGGGGAAACCGGATG 378
 QY 201 ValAlaGlyArgValGluAspValGluLeuValAspLeuAlaAlaAspGlyValGluLeu 220
 Db 377 ATTGCCAGATGTCTCCAGATGTCCAACTGGTAGACTTCGCGGGGATCGAGTAAACTT 318
 QY 221 AlaValGluValLeuAspGlyArgArgValGlyValLeuGluAlaAlaGlyGluGluAla 240
 Db 317 TCGGTGGAAGCTTCAATGTGTAGGGGTAAAGCGTCTCGAAGCGCCCGTGAAGGAGCG 258
 QY 241 ArgHisAspGlyArgLeuAlaHisLeuGlyArgAlaGluAspAspHisAlaIleAlaVal 260
 Db 257 CGACACGATGGCGCTTGGCCCACTTGGATGAGACCGAGATGACCTCTGTAGCAGTT 198
 QY 261 LeuGlyArgAspThrGlnLeuArgValAlaArgAlaHisLeuLeuAspHisAspGlyGln 280
 Db 197 CTTGGCGCGATACCTCAATGTGAGGTGCTTGGCACATCTTGTATCATCGCGGCGAG 138
 QY 281 PheHis 282
 Db 137 TTTTCAT 132
 RESULT 8

PN WO958670-A1.
 XX 18-NOV-1999.
 XX
 XX 07-MAY-1999; 99WO-US10151.
 XX
 XX 08-MAY-1998; 98US-0084842.
 PR 07-OCT-1998; 98US-0103355.
 XX
 PA (CADU-) CADUS PHARM CORP.
 XX
 PI Cismowski M, Duzic E;
 XX
 DR MPI: 2000-072337/06.
 DR P-PSDB; AAY53923.
 XX
 XX A new activator of G protein signalling used to treat disorders
 PT characterized by an aberrant AGS protein activity -
 XX
 PS Example 15; Page 143-144; 162pp; English.
 XX
 CC The present sequence encodes a homologue of a human AGS1 (activator
 CC of G protein signalling (AGS)) protein. The AGS cDNA sequence was
 CC isolated from a human liver cDNA library. The AGS protein exhibits
 CC homology to ras-related G proteins, and contains alterations in
 CC conserved amino acids consistent with a deficiency in GTP hydrolysis
 CC activity. AGS stimulates G protein activity, G protein-coupled signal
 CC transduction and the pheromone response pathway in a receptor-independent
 CC manner. The AGS protein also shows G-gamma selectivity, as measured by
 CC growth assays in yeast expressing various mammalian G-gamma constructs,
 CC and tissue-specific expression, as measured by Northern blot analysis.
 CC The AGS protein can be used to screen for compounds that modulate
 CC cellular signal transduction. The protein is used to treat disorders
 CC characterized by an aberrant AGS protein activity or AGS nucleic acid
 CC expression.
 XX
 SQ Sequence 837 BP; 177 A; 272 C; 233 G; 155 T; 0 other;
 XX
 Alignment Scores:
 Pred. No.: 1,05e-31 Length: 837
 Score: 579.50 Matches: 146
 Percent Similarity: 61.59% Conservative: 24
 Best Local Similarity: 52.90% Mismatches: 93
 Query Match: 41.07% Indels: 13
 DB: 21 Gaps: 5
 US-09-709-103-1f4 (1-282) x AAZ36913 (1-837)
 QY 4 AspAlaAlaLeuLeuValLeu--GlyLeuAlaAlaGlyAlaGlyLeuLeuAlaAspVal 22
 DB 831 GATGATGACACTTCTCCCTGCACGGGCTGGCTTCCGAAAGACCTTGGCTTGATGTA 772
 QY 23 HisGluValAlaValAlaValAlaValAlaValAlaValAlaValAlaValAlaVal 42
 DB 771 CTTGAGGTGACCTGTTACCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTC 712
 QY 43 AlaArgValAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 62
 DB 711 ---CATCTCTTGCACGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 676
 QY 63 LeuLeuValGlnHisValAlaValAlaValAlaValAlaValAlaValAlaValAlaVal 82
 DB 675 -----GTGAAAGCGCTCACCGTACGACGAGATCTTGGATGTCAGGGGGGCTCAT 622
 QY 83 LeuAlaGlyGlnLeuGlnHisGlyGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 102
 DB 621 CTCGTGTGCGACCTTGGCCATCTTAAGAGCACGTAGAACAATCTGCTCCACGTGGTGT 562
 QY 103 LeuLeuGlyArgAspLeuGlnValGlyAlaAlaValAlaValAlaValAlaValAlaVal 122
 DB 561 CTTCTTGGCGCACCTCGAAGTACGCCAGTTCTC---GTGCGCCGACACACACACCTC 505
 QY 123 AspLeuAlaLeuValHisLeuAlaValGlnValAla---ValThrLeuValAlaAlaAsp 141

DB 504 GGCCTCGGTGTGGGACACTGCGCGCACAGCTCGCGCTCTTCTTGTGGACAGAT 445
 QY 142 AspGlnGlyHisValHisValLeuLeuGlyPheValLeuGlnAlaArgLeuGlyValGlu 161
 DB 444 GACCATGGGACAGCTCCGCGCGCTCTTGTGTTCTTCTTACAGCAGACCTTGAACCTCCAG 385
 QY 162 AspLeuLeuProGluProLeuHisLeuLeuGlnGlyValAlaValAlaGlnThrGlnHis 181
 DB 384 GATCGCTTCTGAAGGGCGCTTGACCTCATGGAAGACTCCCGGTTATCCAGCGCTGAAC 325
 QY 182 GlnAspGlnLeuValSerCysGlnAspGlyGlnAlaProHisGlyGlnGlnValAla 201
 DB 324 CAGATGAAGACATCCCTCTGAGATGACAGACCTCCGATGCGGGAAGGGGGT 265
 QY 202 AlaGlyArgValGlnAspValGlnLeuValAspLeuAlaAlaAspGlyValGlnLeuAla 221
 DB 264 GCCAGAGTATCCAGGATGTCGAGCTGTACATGTCGCGCGGATGTTGTATACCTTACG 205
 QY 222 ValGlnValLeuAspGlyArgArgValGlyValLeuGlnAlaAlaGlyGlnGlnAlaArg 241
 DB 204 GTGAAAGTCTTCATGAGTGGGTGTGTACTGTCTTAAGCGGCAATTGAGAGACCGAGA 145
 QY 242 HisAspGlyArgLeuAlaHisLeuGlnGlyArgAlaGlnAspAspHisAlaAlaValLeu 261
 DB 144 CAGATGAGACTCTTCCCGCACCGAGAGGACACCCAGCACCACTACCGGTATGAGATT 85
 QY 262 GlyArgAspThrGlnLeuArgValAlaAlaArgAlaHisLeuLeuAspHis 277
 DB 84 GCGGCGCACACTGAGGTGACATTCCTCCGCTGGAACAAGCTTCATCAT 37
 RESULT 10
 ID ABL92076/c
 XX ABL92076 standard; cDNA; 2832 BP.
 AC ABL92076;
 XX
 DT 30-MAY-2002 (first entry)
 XX
 DE Human Tumour Endothelial Marker polynucleotide SEQ ID NO 174.
 XX
 KW Human; mouse; rat; TEM; tumour endothelial marker; NEM; PEM; cytostatic;
 KW normal endothelial marker; pan-endothelial marker; immunostimulant;
 KW antiangiogenic; tumour; neoangiogenesis; vascularised tumour;
 KW polycystic kidney disease; diabetes; retinopathy; rheumatoid arthritis;
 KW psoriasis; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200210217-A2.
 XX
 PD 07-FEB-2002.
 XX
 PF 01-AUG-2001; 2001WO-US24031.
 XX
 PR 02-AUG-2000; 2000US-222599P.
 PR 11-AUG-2000; 2000US-224360P.
 PR 11-APR-2001; 2001US-282850P.
 XX
 PA (UYGO) UNIV JOHNS HOPKINS.
 XX
 PI St Croix B, Kinzler KW, Vogelstein B;
 XX
 DR MPI: 2002-291856/33.
 XX
 XX An isolated molecule comprising an antibody variable region which
 PT specifically binds to an extracellular domain of a tumor endothelial
 PT marker (TEM) protein, useful for inhibiting tumor growth -
 XX
 PS Disclosure; Page 119-120; 331pp; English.
 XX
 CC The invention relates to an isolated molecule comprising an antibody
 CC variable region which specifically binds to an extracellular domain of a

QY 24Z HIBASPOUYATYUPEWATIAAOWOOCSSJINJN-----

US-09-709-103-1F4 (1-282) x ABL92087 (1-2973)


```

QY 4 AspaAlaAlaLeuValLeu---GlyLeuAlaAlaGlyAlaGlyLeuAlaAspVal 22
Db 1002 GATGGGCACTGTCTCTCCCTCCAGGGCTGGCTTCCCGAGAGACCTTGGCTTGAATGA 943
QY 23 HisGluValAlaValTyrAlaGlyProAlaArgGlyGlyCysHisAspAlaValGlyVal 42
Db 942 CTGGAGGTGACCTGTGGAGCTGGGGCGGCGAGAGGCGAGACCATGCGATAGGGGTG 883
QY 43 AlaArgValAlaAlaAlaAlaAlaAlaGlyProGlnGlnLeuLeuValProGlnArg 62
Db 882 ---CATCTCTTGAAGCGGGCGGCGATGAGAGGGCTGGG----- 847
QY 63 LeuLeuValGlnHisValAlaValLeuHisArgAspLeuAlaValGlnValTyrAlaHis 82
Db 846 -----GTGGAGAGCGGTACCGGTACTGACGAGATCTTGCATGACAGGGCGGGCTCAT 793
QY 83 LeuAlaGlyGlnLeuGlyHisGlyGlyGluArgAlaGlnHisLeuValGlnAlaAlaVal 102
Db 792 CTCTGTGGGCACTGTGGCGATGTAAGAGACCGTACATCTCTGCTTGGTGTG 733
QY 103 LeuLeuGlyArgAspLeuGlyValAlaAlaLeuGlyValAlaHisGlnLeuLeu 122
Db 732 CTCTGTGGGCGACCTCGAAGTGGGCGAGTTCTC---GTCCCGGACACCGACGCTC 676
QY 123 AspLeuAlaLeuValHisLeuAlaValGluValAla---ValThrLeuValAlaAlaAsp 141
Db 675 GGCCTCGGTGGTGGCACTGGGCGGACAGCTCGCGGTGGTCTTCTTGGTGCACAGAT 616
QY 142 AspGlnGlyHisValHisValLeuLeuGlyPheValLeuGlnAlaArgLeuGlyValGlu 161
Db 615 GACCATGGGAGCTCCCGCCCTCTTGGTCTTGTCTTCTTGGAGAGAGTGAACCTCCAG 556
QY 162 AspLeuLeuProGluProLeuHisLeuLeuGlyGlyValAlaValAlaGlnThrGlnHis 181
Db 555 GATCTGCTTGTGAAAGCGCTTGAACCTCATGAGAGGACCTCCGGTTATCCAGGCTGAAC 496
QY 182 GluAspGlyAsnValSerCysGluAspGlyGluAlaProHisGlyGlyGluArgValVal 201
Db 495 CAGAGTGAAGACATCCCTCTGTGAGATGAGACAGCTCCGATGGGGAGAGGGTGGT 436
QY 202 AlaGlyArgValGluAspValGluLeuValAspLeuAlaAlaAspGlyValGluLeuVal 221
Db 435 GCCAGAGGTATCCAGAGTGTGAGCTGTGATGCTGCGCGGAGTGTGTATCTTACG 376
QY 222 ValGluValLeuAspGlyArgValGlyValLeuGlnAlaAlaGlyGluGlnValArg 241
Db 375 GTGGAGTCTCTCATGTGTGAGTGTGATGCTCAAAAGCGGCATGAGAGCGAGA 316
QY 242 HisAspGlyArgLeuAlaHisLeuGlyArgAlaGluAspAspHisAlaAlaValLeu 261
Db 315 CAGCATGAGAGCTTGTCCACCCGAGAGGACACCCAGACACCACTGCGTATGAGTTT 256
QY 262 GlyArgAspThrGlnLeuArgValAlaArgAlaHisLeuLeuAspHis 277
Db 255 GGGGGGACACATGAGGTCAGTCTCCCGGTGACAAAGTCTTCAAT 208

RESULT 12
ABK71563/C
ID ABK71563 standard, cDNA, 3427 BP.
XX
XX ABK71563;
XX
XX 30-JUL-2002 (first entry)
XX
XX Human dithp polynucleotide #29.
XX
XX Human, dithp, diagnostic and therapeutic polynucleotide; gene; ss; bone;
XX cell proliferative disorder; cancer; tumour; autoimmune disorder; brain;
XX inflammatory disorder; viral infection; bacterial infection; seizure;
XX fungal infection; parasitic infections; neurological disorder; breast;
XX endocrine disorder; metabolic disorder; developmental disorder; cervix;
XX gastrointestinal disorder; transport disorder; gene therapy; kidney;

```

```

KM adrenal gland; bone marrow; lung; ovary; pancreas; prostate; spleen;
KM skin; testis; thymus.
XX
XX Homo sapiens.
OS
XX WO200220754-A2.
XX
XX 14-MAR-2002.
XX
XX 29-AUG-2001; 2001WO-US27127.
XX
XX 05-SEP-2000; 2000US-229747P.
XX 05-SEP-2000; 2000US-229748P.
XX 05-SEP-2000; 2000US-229749P.
XX 05-SEP-2000; 2000US-229750P.
XX 05-SEP-2000; 2000US-229751P.
XX 05-SEP-2000; 2000US-230533P.
XX 05-SEP-2000; 2000US-230505P.
XX 05-SEP-2000; 2000US-230514P.
XX 05-SEP-2000; 2000US-230515P.
XX 05-SEP-2000; 2000US-230517P.
XX 05-SEP-2000; 2000US-230518P.
XX 05-SEP-2000; 2000US-230519P.
XX 05-SEP-2000; 2000US-230597P.
XX 05-SEP-2000; 2000US-230598P.
XX 05-SEP-2000; 2000US-230599P.
XX 05-SEP-2000; 2000US-230610P.
XX 05-SEP-2000; 2000US-230865P.
XX 05-SEP-2000; 2000US-230988P.
XX 07-SEP-2000; 2000US-230951P.
XX 07-SEP-2000; 2000US-231163P.
XX 07-SEP-2000; 2000US-231167P.
XX
XX (INCYTE- ) INCYTE GENOMICS INC.
XX
XX Stuart J, Lincoln SE, Altus CM, Dufour GE, Chalup MS, Hillman DJ,
XX Jones AL, Yu JY, Wright RJ, Gietzen D, Liu TF, Yap PE, Dahl CR,
XX Moriyma MG, Bradley DJ, Rohatgi SD, Harris B, Roseberry AM,
XX Gerstein EH, Peralta CH, David MH, Panzer SR, Flores V, Dafo A,
XX Marwaha K, Chen AJ, Chang SC, Au AP, Inman RR;
XX WPI; 2002-383054/41.
XX P-PSDB; ABG59971.
XX
XX An isolated polynucleotide useful in diagnostics and therapeutics -
XX
XX Claim 1; Page 419-420; 686pp; English.
XX
XX The invention relates to human diagnostic and therapeutic (dithp)
XX polynucleotides and their associated polypeptides (Dithp polypeptides).
XX The sequences of the invention are used in the treatment and diagnosis of
XX cell proliferative disorders (e.g. atherosclerosis, cirrhosis), cancers
XX (e.g. tumours of the adrenal gland, bone, bone marrow, brain, breast,
XX cervix, kidney, lung, ovary, pancreas, prostate, skin, spleen, testis or
XX thymus), autoimmune/inflammatory disorders (e.g. asthma, bronchitis,
XX psoriasis, osteoporosis), viral infections, bacterial infections, fungal
XX infections, parasitic infections, developmental disorders (e.g. anaemia,
XX epilepsy), seizure disorders (e.g. cerebral palsy, spina bifida),
XX endocrine disorders (e.g. diabetes), neurological disorders (e.g. stroke,
XX e.g. obesity, diabetes), neurodegenerative disorders (e.g. Alzheimer's
XX amyotrophic lateral sclerosis, multiple sclerosis), gastrointestinal
XX disorders (e.g. ulcerative colitis, Crohn's disease), and transport disorders
XX (e.g. myotonic dystrophy, cataracts, peripheral neuropathy). Sequences
XX ABK71563-ABK71809 represent human dithp polynucleotides of the invention.
XX
XX Sequence 3427 BP; 682 A; 1145 C; 935 G; 665 T; 0 other.
XX
XX Alignment Scores:
XX Pred. No.: 4,296-31
XX Score: 579.50
XX Percent Similarity: 61.59%
XX Best Local Similarity: 52.90%
XX Length: 3427
XX Matches: 146
XX Conservative: 24
XX Mismatches: 93

```

Tue Dec 31 15:27:52 2002

Query Match: 41.07% Indels: 13
DB: 24 Gaps: 5

US-09-709-103-1F4 (1-282) x ABK71563 (1-3427)

Qy 4 AspalalaLeuValLeu---GlyLeuAlaAlaGlyAlaGlyLeuLeuAlaAspVal 22
Db 1438 GATGGTGCACCTTGTCCCTCTCACGGGCTGGCTCCCGAAGGACCTTGGCTTGTATGTA 1379

Qy 23 HisGluValAlaValTyrAlaGlyProAlaArgGluGlyCysHisAspAlaLysGlyVal 42
Db 1378 CTTGAGGTCACTTGTACGCTGGCGGGCGGCGAAGGGCGAGACCATGCGCATAGGCGTC 1319

Qy 43 AlaArgValAlaAlaAlaAlaAlaGlyProGlnGlnLeuLeuValProGlnArg 62
Db 1318 ---CATCTCTTACGCGCGCGCATGCAGAGGGCTTGGG----- 1283

Qy 63 LeuLeuValGlnHisValAlaValLeuHisArgAspLeuAlaValGlnValTrpAlaHis 82
Db 1282 -----GTGGAAGGCGTCACTGCTACCGAGATCTTGCATGCGAGGGCGGCTCAT 1229

Qy 83 LeuAlaGlyGlnLeuGlyHisGlyGluGluArgAlaGluHisLeuValGlnAlaVal 102
Db 1228 CTCGTGTGGCAGCTTGGCCATCTGTAAGAGCACGTAGAACATCTCGTCCACGTTGGTGT 1169

Qy 103 LeuLeuGlyArgAspLeuValGlyAlaAlaLeuGlyValValAlaHisGlnLeuLeu 122
Db 1168 CTTCTTGGCGGACACCTCGAAGTAGGCGCATCTTC---GTGCGCGCACACCGAGCTC 1112

Qy 123 AspLeuAlaLeuValHisLeuAlaValGluValAla---ValThrLeuValAlaAlaAsp 141
Db 1111 GGCCTCGTGGTGGGCACCTGGCGGCACAGCTCGCGTGTCTTGTTCGCCACAGAT 1052

Qy 142 AspGlnGlyHisValHisValLeuLeuGlyPheValLeuGluAlaArgLeuGlyValGlu 161
Db 1051 GACCATGGCGAGCTCCCGCGCTCTCTTGTCTTTCAGCGAGGACTTACCTCCAG 992

Qy 162 AspLeuLeuProGlnProLeuHisLeuLeuGluGlyValAlaValValGlnThrGluHis 181
Db 991 GATCTGCTTCTGAAGGCGCTTGACCTCATCGAAGGACTCCCGTTCAGGCTGAACAC 932

Qy 182 GlnAspGluAsnValSerCysGluAspGlyGluAlaProHisGlyGlyGluArgValVal 201
Db 931 CAGGATGAACACATCCCTGTAGGATGGACACCTCGCATGGCGGGAAGGGTGGT 872

Qy 202 AlaGlyArgValGluAspValGluLeuValAspLeuAlaAlaAspGlyValGluLeuAla 221
Db 871 GCCAGAGGTATCCAGGATGTCAGCTGGTACATGTGCGCGGATGTTGTATACCTTACG 812

Qy 222 ValGluValLeuAspGlyArgArgValGlyValLeuGluAlaAlaGlyGluGluAlaArg 241
Db 811 GTGGAAGTCTCGATGTGGTGTGTACTGGTCTCTCAAGCGGCCATTTGAGGAAGCGAGA 752

Qy 242 HisAspGlyArgLeuAlaHisLeuGlyArgAlaGluAspHisAlaIleAlaValLeu 261
Db 751 CACGATGGAGCTCTTGGCCACCCGAGAGGACCCAGCACCCACCATGGGTATGATGTTT 692

Qy 262 GlyArgAspThrGlnLeuArgValAlaArgAlaHisLeuLeuAspHis 277
Db 691 GCGCGGACACTGAGCGTGCAGTCCCGCTGCACAAAGTCTTCATCAT 644

RESULT 13
ABL92134/c
ID ABL92134 standard; cDNA; 3020 BP.
XX
AC ABL92134;
XX
XX
DT 30-MAY-2002 (first entry)
XX
DE Mouse Tumour Endothelial Marker polynucleotide SEQ ID NO 292.
XX
KW Human; mouse; rat; TEM; tumour endothelial marker; NEM; PEM; cytostatic;
KW normal endothelial marker; pan-endothelial marker; immunostimulant;

KW antiangiogenic; tumour; neoangiogenesis; vascularised tumour;
KW polycystic kidney disease; diabetes; retinopathy; rheumatoid arthritis;
KW psoriasis; gene; ss.
XX Mus musculus.
XX WO200210217-A2.
XX PD 07-FEB-2002.
XX PF 01-AUG-2001; 2001WO-US24031.
XX PR 02-AUG-2000; 2000US-222599P.
XX PR 11-AUG-2000; 2000US-224360P.
XX PR 11-APR-2001; 2001US-282850P.
XX (UYJO) UNIV JOHNS HOPKINS.
XX St Croix B, Kinzler KW, Vogelstein B;
XX WPI; 2002-291856/33.
XX P-PSDB; ABB90781.
XX An isolated molecule comprising an antibody variable region which specifically binds to an extracellular domain of a tumor endothelial marker (TEM) protein, useful for inhibiting tumor growth -
XX Disclosure; Page 294-295; 33pp; English.
XX The invention relates to an isolated molecule comprising an antibody variable region which specifically binds to an extracellular domain of a tumor endothelial marker (TEM) protein selected from ABB90732, ABB90740, ABB90749, ABB90750 and ABB90769. The antibodies which bind to TEM proteins have cytostatic, immunostimulant and antiangiogenic activity. They are useful for inhibiting tumor growth, neoangiogenesis in subjects bearing a vascularised tumour, polycystic kidney disease, diabetic retinopathy, rheumatoid arthritis and psoriasis. Human, mouse and rat TEM genes and the encoded proteins (ABL92075-ABL92141 and ABB90721-ABB90789) are disclosed, as are marker oligonucleotide sequences: tumour endothelial markers (TEM) ABL91996-ABL92041 and ABL92143-ABL92191; normal endothelial markers (NEM) ABL92042-ABL92074; and pan-endothelial markers (PEM) ABL91903-ABL91995.
XX Sequence 3020 BP; 690 A; 848 C; 780 G; 702 T; 0 other;
Alignment Scores:
Pred. No.: 1.45e-25 Length: 3020
Score: 498.50 Matches: 127
Percent Similarity: 58.42% Conservative: 36
Best Local Similarity: 45.52% Mismatches: 103
Query Match: 35.33% Indels: 13
DB: 24 Gaps: 5
US-09-709-103-1F4 (1-282) x ABL92134 (1-3020)
Qy 1 LeuAlaAspAspAlaAlaLeuLeuValLeu---GlyLeuAlaAlaGlyAlaGlyLeuLeu 19
Db 1179 CTTCTCACTGGATGCTACACTTGTCCCTCTCGGGCTTGGCCCTCCGCTAGGACCTTGGC 1120
Qy 20 AlaAspValHisGluValAlaValTyrAlaGlyProAlaArgGluGlyCysHisAspAla 39
Db 1119 CTTGATGTACTTGTAGTCACTGTTGACACTGGGCGCTCGGCAAGGTTGAGACCATGCC 1060
Qy 40 LysGlyValAlaAlaArgValAlaAlaAlaAlaAlaAlaGlyProGlnGlnLeuVal 59
Db 1059 ATAGGACCTCGACCTT---AGTCGCGCATGTCAGAGGCGCGGG----- 1015
Qy 60 ProGlnArgLeuLeuValGlnHisValAlaValLeuHisArgAspLeuAlaValGlnVal 79
Db 1014 -----GTGAAAAGCATCGCGTACTGTCACCGAGATCTTATGTGTGAGTGC 970
Qy 80 TrpAlaHisLeuAlaGlyGlnLeuGlyHisGlyGluGluArgAlaGluHisLeuValGln 99

Db 969 AGGGCTCATCTATGGGCGACCTTGGCCATGCTGAACAGCAGCATGATCTCTTAC 910
 QY 100 AAlaAlaValLeuLeuGlyArgAspLeuGluValGlyAlaAlaLeuGlyValAlaHis 119
 Db 909 ATTAGTCTCTTCTGGTGAACCTCGAATAGCGCAGTTTC---ATACACAGAC 853
 QY 120 GlnLeuLeuAspLeuAlaLeuValHisLeuAlaValGluValAlaValThr---LeuVal 138
 Db 852 CAGCAGCTCAGCTCCATGAGCAGGAGCCTGGCGCAGCAGCTCAGCTGGTCAATCTTCTT 793
 QY 139 AAlaAlaAspAspGlnGlyHisValHisValLeuLeuGlyPheValLeuGluValArgLeu 158
 Db 792 CCCACAGATACCCATGAGCAGCTGCTGCTCTCTTGGTTTATCTTTCAGGACAGACTT 733
 QY 159 GlyValGluAspLeuLeuProGlnProLeuHisLeuLeuGluGlyValAlaValAlaGln 178
 Db 732 GACCTCCAGGATCTGCTTCTGAGGCGCTTGAACCTCATCAAGAGATCCCGGCTATCCAG 673
 QY 179 ThrGluHisGlnAspGluAsnValSerCysGluAspGlyGluValAlaProHisGlyGln 198
 Db 672 GCTGACACCCAGATGAAGACATCTCTGTAGAGATGAGAGCCGCGCATGCGCAGGAA 613
 QY 199 ArgValValAlaGlyArgValGluValGluValGluValAlaAspGluValAla 218
 Db 612 TGGGTGCTTCCAGAGGTGTCCAGATATCCAGCTGTATCATGTCTCCCGGATGTTGA 553
 QY 219 GluLeuAlaValGluValLeuAspGlyArgArgValGlyValLeuGluValAlaGlyGln 238
 Db 552 CACCTTGGATGAAGATCTCTGATATGAGGCGGTGTACTGCTCTCAAGAGCGCATTTAG 493
 QY 239 GluAlaArgHisAspGlyArgLeuAlaHisLeuGlyValArgAlaGluAspHisAlaIle 258
 Db 492 GAACCGGAGACATGAGCTCTTGGCCACTCGGAGGACCCAGAGCAGCAGCATGCGGTA 433
 QY 259 AlaValLeuGlyArgAspThrGlnLeuArgValAlaArgAlaHisLeuLeuAspHis 277
 Db 432 GAGCTCTTACGAGGACCATTTAGTGTGACGTTCCACACTGGAAGAGCTTATCAT 376
 RESULT 14
 ID AAS90571/c
 XX AAS90571 standard; cDNA, 951 BP.
 AC AAS90571;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE DNA encoding novel human diagnostic protein #26375.
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US08631.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX
 PA (HYSE-) HYSEQ INC.
 PI Drmanac RT, Liu C, Tang YT;
 XX
 DR MPI: 2001-639362/73.
 DR P-PSDB; ABG26384.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensic, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.

XX
 PS Claim 1: SEQ ID No 26375; 103bp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human
 CC diagnostic coding sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 951 BP; 202 A; 291 C; 293 G; 165 T; 0 other;
 Alignment Scores:
 Pred. No: 1.1e-24 Length: 951
 Score: 478.50 Matches: 133
 Percent Similarity: 57.85% Conservative: 18
 Best Local Similarity: 50.96% Mismatch: 87
 Query Match: 33.91% Indels: 23
 DB: 23 Gaps: 8
 US-09-709-103-1f4 (1-282) x AAS90571 (1-951)
 QY 4 AspAlaAlaLeuLeuValLeu---GlyLeuAlaAlaGlyValGlyLeuLeuAlaAspVal 22
 Db 876 GATGTCGACATTTCTCCCTTCAAGGCTTCCGGAAGACCTTGGCTGATGTA 817
 QY 23 HisGluValAlaValTyrAlaGlyProAlaArgGluGlyCysHisAspAlaGlyVal 42
 Db 816 CTGAGGTCACTGTGAGCTGTGGCGGCGGAGGCGAGACCATGATGCGCTC 757
 QY 43 AlaArgValAlaAlaAlaAlaAlaAlaGlyProGlnGlnLeuLeuValProGlnArg 62
 Db 756 ---CATCTCTTGAAGCGCGCATGACAGAGGCGCTGG----- 721
 QY 63 LeuLeuValGlnHisValAlaValLeuHisArgAspLeuAlaValGlnValTyrAlaHis 82
 Db 720 -----GTGAAGGCGTCAACCTGATCGACGAGATCTTGGATGCGAGGCGGCTCAT 667
 QY 83 LeuAlaGlyGlnLeuGlyHisGlyGlnGluArgAlaGlnHisLeuValGlnAlaValAla 102
 Db 666 CTCGTGTGCGACGCTTGGCATGTAAGAGCAGTGAACATCTCGTCCACGTTGGTGT 607
 QY 103 LeuLeuGlyArgAspLeuGlyValGlyAlaAlaLeuGlyValValAlaHisGlnLeuLeu 122
 Db 606 CTTCTTGGCCAGACCTTGAAGTGAAGGCGAGTTTCTC---GTGCGCCAGACACGAGCTC 550
 QY 123 AspLeuAlaLeuValHisLeuAlaValGluValAla---ValThrLeuValAlaAlaAsp 141
 Db 549 GGCCTCGGTGGTGGGACCTGCGGACAGTCCCGTGTCTTGTGGCAGACAT 490
 QY 142 AspGlnGlyHisValHisValLeuLeuGlyPheValLeuGluValArgLeuGlyValGln 161
 Db 489 GACCATGGGACCTCCCGCGCTCTTGGTTCTTCTGAGCAGAGATGATGACCTCCAG 430
 QY 162 AspLeuLeu-----ProGlnProLeuHisLeuLeuGluGlyValAlaVal 176
 Db 429 GATCTGCTAAGTCCAGACACATCTCCGACCCA-----TGCCTTGCCCTGCGGCC 379

types and for investigating cell differentiation. The method allows the methylation status of many C residues to be determined simultaneously. AB013410-AB054121 represent genomic DNA sequences used to illustrate the method for determining the degree of cytosine methylation described in the disclosure of the invention.

XX
Sequence 624 BP: 127 A; 68 C; 206 G; 223 T; 0 other;

Alignment Scores:		
Pred. No.:	5.65e-21	624
Score:	422.00	81
Percent similarity:	80.15%	28
Best Local Similarity:	59.56%	Conservative:
Query Match:	29.91%	Mismatches: 27
DR:	24	Indels: 1
		Gaps: 0

US-09-709-103-1F4 (1-282) X ABO28318 (1-624)

1	LeuAlaAspAspAlaLeuLeuValLeuGlyLeuAlaAlaGlyAlaGlyLeuLeuAla	20
QY		
409	CTAACTAATACGCAACGGCTCTCTTATCTTAACCTTACCAGCGCTAACCTTCTCGCG	350
Db		
21	AspValHisGluValAlaValTyrAlaGlyProAlaArgGluGlyCysHisAspAlaLys	40
QY		
349	AAATATACATAAATTCGTATATACGCTAAACCGCGCGCGAAATAATCACGNATCCAAA	290
Db		
41	GlyValAlaArgValAlaAlaAlaAlaAlaAlaGlyProGlnGlnLeuLeuValPro	60
QY		
289	AA-ATCGCCGGAATCGCGCGCGCGCGCGCTACCGACCGCAACAACCTCTTATTTCGG	231
Db		
61	GlnArgLeuLeuValGlnHisValAlaValLeuHisArgAspLeuAlaValGlnValTrp	80
QY		
230	CAACGGCTTCTTATACAAACACGTCGCAATACTTACACCGAAACCTTACGATACAAATCTAA	171
Db		
81	AlaHisLeuAlaGlyGlnLeuGlyHisGlyGluGluArgAlaGluHisLeuValGlnAla	100
QY		
170	ACTCATCTCGCTAAACAACCTTAAACCATATACGAAACCGCGGAACATCTATATCCAACT	111
Db		
101	AlaValLeuLeuGlyArgAspLeuGluValGlyAlaAlaLeuGlyValValAlaHisGln	120
QY		
110	ACTATTCTTCTTAACCGGAATCTTCGAATAATACGGAACGCTAAAANTCGTCGCCCAACCA	51
Db		
121	LeuLeuAspLeuAlaLeuValHisLeuAlaValGluValAlaValThr	136
QY		
50	CTATCTGATCTTCGGCTTATCTCACTTCGGGATAAAAAATTCGGCATCACC	3
Db		

Search completed: December 30, 2002, 16:35:28
Job time : 375 secs

[illegible]

RESULT 15

ABQ28318/c
ID ABO28318 standard; DNA; 624 BP.

XX ABO28318:

XX
DT 12-JUL-2002 (first entry)

XX
01-chemical code for detecting cytosine methylation SEQ ID NO 14909.

XX

XX	Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
KW	drug; side effect; cancer; central nervous system; cardiovascular;
KW	gastrointestinal; respiratory system; single nucleotide polymorphism;
KW	SNP; cell differentiation; ds.

AA Homo sapiens.

XX PN WO200218632-A2.

XX
07-MAR-2002.XX
DE 01-SEP-2001: 2001WO-EP10074.

XX
01-SEP-2000: 2000DE-1043826.

PR 05-SEP-2000; 2000DE-1044543.
vv

PA (EPIG-) EPIGENOMICS AG.

XX PI Olek A Pienenbrock C, Berlin K, Guetig D;

XX
WPB: 2002-371829/40

XX
PT
PT
PT
PT

Determining the degree of cytosine methylation in genomic DNA, useful for diagnosis and prognosis, comprises selective hybridization of analices from chemically treated DNA -

XX
ns
XX
ns
claim 12. 56pp + Sequence Listing; 56pp; German.

This invention describes a novel method for determining the degree of methylation of a particular cytosine in a motif 5'-CpG-3', present in a genomic sample of DNA. The sample is treated chemically to convert cytosine (C) but not methylated C, to uracil, then part of the genomic DNA that contains the target C is amplified to form a labeled amplicon. The amplicon is hybridised to two classes, each with at least one member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the degree of hybridisation to both classes is determined from the label on the amplicon. From the ratio of labels hybridised to the two classes of oligomers, the degree of methylation is calculated. The method is used: (i) for diagnosis and/or prognosis of side effects of therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders of the central nervous, cardiovascular, gastrointestinal and respiratory systems etc., particularly by detecting mutations or single nucleotide polymorphisms (SNP/s); and (ii) for differentiation of cell or tissue

Gencore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: December 30, 2002, 14:47:25 ; Search time 3135 Seconds

(without alignments)
2617.860 Million cell updates/sec

Title: US-09-709-103-1f4
Perfect score: 1411
Sequence: 1 LADDAALVIGLAAGGLA.....RDTQRAVARAHLLDRGQFH 282

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xlp
-Q/cgm2.1/USPTO.spool/SULIVAN14/runat.30122002.144046.21500/app.query.fasta_1.455
-DB=Gembd1 -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPC=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blsuns62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pro -NORM=ext -HEADSITE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=SULIVAN14@cgm.1.1.3637 @runat.30122002.144046.21500 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MMAB -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
Gembd1: *
1: gb_ba: *
2: gb_htg: *
3: gb_in: *
4: gb_om: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vl: *
15: em_ba: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_mu: *
20: em_or: *
21: em_ov: *
22: em_pat: *
23: em_pl: *
24: em_pi: *
25: em_pi: *
26: em_ro: *
27: em_sts: *
28: em_un: *

Seq 1 Frame 4

29: em_vi: *
30: em_htg_hum: *
31: em_htg_inv: *
32: em_htg_other: *
33: em_htg_mus: *
34: em_htg_pln: *
35: em_htg_rtd: *
36: em_htg_mam: *
37: em_htg_vrt: *
38: em_ey: *
39: em_htgo_hum: *
40: em_htgo_mus: *
41: em_htgo_other: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1411	100.0	846	AF498923	AF498923 Homo sapi
2	1411	100.0	979	AF172846	AF172846 Homo sapi
3	1411	100.0	1740	AF069506	AF069506 Homo sapi
4	1411	100.0	1758	BC018041	BC018041 Homo sapi
5	1398	99.1	1746	AF177335	AF177335 Homo sapi
6	1392	98.7	1187	AF153192	AF153192 Homo sapi
7	1348	95.5	4990	AF222979	AF222979 Homo sapi
8	1338	95.5	18334	AC020558	AC020558 Homo sapi
9	1318	93.4	58882	AC073621	AC073621 Homo sapi
10	1314	93.1	183598	AC090608	AC090608 Homo sapi
11	963.5	68.3	1616	AF239157	AF239157 Rattus no
12	963.5	67.3	1612	BC034166	BC034166 Mus muscu
13	949.5	66.8	1621	AF009246	AF009246 Mus muscu
14	942.5	64.8	179124	AL603710	AL603710 Mouse DNA
15	914	64.8	247899	AC025909	AC025909 Mus muscu
16	914	45.4	162504	AC122995	AC122995 Rattus no
17	641	41.1	2832	AX393244	AX393244 Sequence
18	579.5	41.1	2832	AF279143	AF279143 Homo sapi
19	579.5	41.1	2973	AX393267	AX393267 Sequence
20	579.5	41.1	3058	BC013419	BC013419 Homo sapi
21	547	38.8	2699	HSMB03172	HSMB03172 Homo sapi
22	547	35.3	3020	AX393352	AX393352 Sequence
23	498.5	32.4	3469	AF134409	AF134409 Rattus no
24	456.5	26.3	114771	HS569D19	HS569D19 Human DNA
25	371.5	23.4	211071	AC076974	AC076974 Mus muscu
26	329.5	17.7	3391	BC030660	BC030660 Homo sapi
27	249.5	17.5	597	AY056037	AY056037 Homo sapi
28	247.5	17.5	597	AY059641	AY059641 Homo sapi
29	247.5	17.5	597	AB076888	AB076888 Homo sapi
30	247.5	17.5	2827	AK096600	AK096600 Homo sapi
31	247.5	17.3	177540	AC006538	AC006538 Homo sapi
32	247.5	16.2	1249	AX430418	AX430418 Sequence
33	244.5	16.0	66031	AC118046	AC118046 Mus muscu
34	228.5	15.8	153386	AC103000	AC103000 Rattus no
35	222.5	15.4	2294	BC026377	BC026377 Mus muscu
36	217.5	14.3	3670	AY129454	AY129454 Drosophila
37	197.5	14.0	126901	AC091518	AC091518 Mus muscu
38	197.5	14.0	201320	AC073816	AC073816 Mus muscu
39	197.5	13.9	597	AY059640	AY059640 Mus muscu
40	196.5	13.5	2145	HSMB03027	HSMB03027 Homo sapi
41	190	13.4	1913	BC012362	BC012362 Homo sapi
42	188.5	13.4	140794	AC069415	AC069415 Homo sapi
43	188.5	13.4	165733	AC069072	AC069072 Homo sapi
44	188.5				
45	188.5				

RESULT 1

ALIGNMENTS

QY	121	LeuLeuAspLeuAlaLeuValHisLeuAlaValGluValAlaValThrLeuValAlaAla	141
Db	486	CTGCTCGATCTCGCGTGTGTCCACCTCGCGGTAGAAAGTCGCGGTCACTCTGTTCGCCGA	427
QY	141	AspAspGlnGlyHisValHisValLeuLeuGlyPheValLeuGluAlaArgLeuGlyVal	160
Db	426	GATGACCAAGGGGACGTCACGTTCTCCTTGGTTTGTCTTGAGCAAGACTTGGTGTC	367
QY	161	GluAspLeuLeuProGluProLeuHisLeuLeuGluGlyValAlaValValGlnThrGlu	180
Db	366	GAGGATCTGCTGCCTGAGCGCGTGCACCTCTCGAAGGAGTCGCGGTTGTCCAGACTGAA	307
QY	181	HisGlnAspGluAsnValSerCysGluAspGlyGluAlaProHisGlyGlyGluArgVal	200
Db	306	CACCCAGATGAACACGTCCTCTGTAGGATGGAGAGCCCGCATGGCGGGAAACGGGTG	247
QY	201	ValAlaGlyArgValGluAspValGluLeuValAspLeuAlaAlaAspGlyValGluLeu	220
Db	246	GTTCGCCGACGTCGAGGATGTCGAGCTGTTAGACCTCGCCGCGATGCAGTAGAACCT	187
QY	221	AlaValGluValLeuAspGlyArgArgValGlyValLeuGluAlaAlaGlyGluGluAla	240
Db	186	GCGGTGGAAATCTTCGATGGTAGCGGTAGCGCTCTCGAAGCGCGCGGTGAGGAAGCG	127
QY	241	ArgHisAspGlyArgLeuAlaHisLeuGlyArgAlaGluAspAspHisAlaIleAlaVal	260
Db	126	CGACACAGATGGCCGCTCTTCCACCTTGGACGAGCCGAGGATGACCATCGATAGCAGTT	67
QY	261	LeuGlyArgAspThrGlnLeuArgValAlaAlaArgAlaHisLeuLeuAspHisArgGlyGln	280
Db	66	CTTTGGCGGGATACTCAGTCCGAGTCGCTCGGGCACATCTTTGATCATCGCGGCCAG	7
QY	281	PheHis 282	
Db	6	TTTCAT 1	
RESULT 2			
AF172846/c			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
gene			
CDS			

```

/protein_id="AA01364.1"
/db_xref="GI:6014489"
/translation="MKLAAMIKMCPDSELSIPAKNCYRWILSSKYKTAIVSRF
LTERFEDATPTIEDFHRKFTYSGEYVOLDLIDTSGNHPPAMRRLSLTGDVFLV
FSLDNDSFEVORLEQOILDTKSKLNKTKENVDPVPIVCGNKGDPRDREVDREI
EQLVDDPORCAVFEISAKNSLSDMPALPAMAKLPEMSPDHARKVSVQCVLH
KVALRKKLIRAGSGGGGSDPDARFIVAPFARRPSVHSDLMYIREKASAGQADKE
RCVTS"
BASE COUNT      195 a      347 c      284 g      153 t
ORIGIN

```

Alignment Scores:

```

Pred. No.:      1,36e-58      Length:      979
Score:          1411.00      Matches:      282
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    100.00%      Indels:      0
DB:             Gaps:      0

```

US-09-709-103-1f4 (1-282) x AF172846 (1-979)

```

QY      1 LeuAlAspAspAlaAlaLeuValLeuGlyLeuAlaAlaGlyLeuLeuAla 20
      |||
DB      909 CTAGCTGATGACCGACGCGCTCTTGCTCTTGCGCTGCGCTGCGCTGCGCTGCGG 850
QY      21 AspValHisGlyValAlaValAlaGlyProAlaArgGlyGlyCysHisAspAlaLys 40
      |||
DB      849 GATGTACATGAGGTCTGCTGTGAGCTGGGCGCGCGCGCAAGGTGTGCCACGATGCCA 790
QY      41 GlyValAlaArgValAlaAlaAlaAlaAlaAlaGlyProGlnLeuLeuValPro 60
      |||
DB      789 GGGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 730
QY      61 GlnArgLeuLeuValGlnHisValAlaValLeuHisArgAspLeuAlaValGlnValTrp 80
      |||
DB      729 CACGCGCTCTTGCTGACGACGCTGCGACGACGACGACGACGACGACGACGACGACGCTG 670
QY      81 AlaHisLeuAlaGlyGlnLeuGlyHisGlyGlyGlnArgAlaGlnHisLeuValGlnAla 100
      |||
DB      669 GCTCATCTGCTGCGCGACGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 610
QY      101 AlaValLeuLeuGlyArgAspLeuGlyValGlyAlaAlaLeuGlyValValAlaHisGln 120
      |||
DB      609 GCTGCTCTCTTGCGCGCGAGCTCTGAGAGTGGCGCGCGCGCTGCGCGCGCGCGCGCG 550
QY      121 LeuLeuAspLeuAlaLeuValHisLeuAlaValAlaGlnValAlaThrLeuValAlaAla 140
      |||
DB      549 CTGCTCGATCTGCGCGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 490
QY      141 AspArgGlnGlyHisValHisValLeuLeuGlyPheValLeuGlnAlaArgLeuGlyVal 160
      |||
DB      489 GATGACCAAGGGGACGCTCCACGCTTCTCTGTTTGTCTTGAGGCAAGACTGGTGTCTC 430
QY      161 GlnAspLeuLeuProGlnProLeuHisLeuLeuGlnGlyValAlaValAlaGlnThrGln 180
      |||
DB      429 GAGGATCTCTGCTCTTACGCGCTGCGCTCTCTGAGGAGAGTGGCGGTGTCTCAGACTGAA 370
QY      181 HisGlnAspGlnAsnValSerCysGlnAspGlyGlnAlaProHisGlyGlyLysArgVal 200
      |||
DB      369 CACCAAGATGAAGAAAGCTCTCTGTAGATGAGAGAGCGCGCGCGCGCGCGCGCGCGCG 310
QY      201 ValAlaGlyArgValGlnAspValGlnLeuValAspLeuAlaAlaAspGlyValGlnLeu 220
      |||
DB      309 GTTGGCGGACGCTGCAAGATGCTGAGCTGTAACCTTCCGCGCGCGATGAGTAAGACTT 250
QY      221 AlaValGlnValLeuAspGlyArgArgValGlnValLeuGlnAlaAlaGlyGlnGlnAla 240
      |||
DB      249 GCGGTGAAGTCTCGATGTAGGCTGTAGGCGGTCTCTGAAAGCGCGCGGTAGGAGAGG 190
QY      241 ArgHisAspGlyArgLeuAlaHisLeuGlyArgAlaGlnAspAspHisAlaAlaVal 260
      |||
DB      189 CGACAGATGCGCGCTCTTCCCGCCTTGACGAGCGGAGAGATGACCATCGCATACAGTT 130

```

```

QY      261 LeuGlyArgAspThrGlnLeuArgValAlaArgAlaHisLeuLeuAspHisArgGlyGln 280
      |||
DB      129 CTGTCGCGGATCTACCTCGAGTGGCTGCGGCGACATCTTCTTATCATCTGCGGCGAG 70
QY      281 PheHis 282
      |||
DB      69 TTTCAAT 64

```

RESULT 3

AF069506/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PubMed

REFERENCE

AUTHORS

TITLE

JOURNAL

Submitted (01-JUN-1998)

Pharmaceutical Corporation, 777 Old Saw Mill River Rd., Tarrytown,

NY 10591, USA

FEATURES

source

1. 1740

/organism="Homo sapiens"

/db_xref="taxon:9606"

/tissue_type="liver"

/dev_stage="adult"

1. 1740

/gene="AGS1"

146. .991

/note="ras-related G protein, similar to Mus musculus

product encoded by GenBank Accession Number AF009246"

/product="activator of G protein signaling"

/codon_start=1

/protein_id="AAD34206.1"

/db_xref="GI:4959038"

/translation="MKLAAMIKMCPDSELSIPAKNCYRWILSSKYKTAIVSRF

LTERFEDATPTIEDFHRKFTYSGEYVOLDLIDTSGNHPPAMRRLSLTGDVFLV

FSLDNDSFEVORLEQOILDTKSKLNKTKENVDPVPIVCGNKGDPRDREVDREI

EQLVDDPORCAVFEISAKNSLSDMPALPAMAKLPEMSPDHARKVSVQCVLH

KVALRKKLIRAGSGGGGSDPDARFIVAPFARRPSVHSDLMYIREKASAGQADKE

RCVTS"

BASE COUNT

ORIGIN

422 a

546 c

483 g

289 t

Alignment Scores:

```

Pred. No.:      2.31e-58      Length:      1740
Score:          1411.00      Matches:      282
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    100.00%      Indels:      0
DB:             Gaps:      0

```

US-09-709-103-1f4 (1-282) x AF069506 (1-1740)

```

QY      1 LeuAlAspAspAlaAlaLeuValLeuGlyLeuAlaAlaGlyLeuLeuAla 20
      |||
DB      991 CTAGCTGATGACCGACGCGCTCTTGCTCTTGCGCTGCGCTGCGCTGCGCTGCGG 932

```


Qy 21 AspValHisGluValAlaValTyAlaGlyProAlaArgGluGlyCysHisAspAlaLys 40
 Db 931 GATGTACATGAGTTCGCTGTACCTGGCGCGCGCGGAGAGGTGCCAGTGCACAA 872
 Qy 41 GlyValAlaArgValAlaAlaAlaAlaAlaAlaGlyProGlnGlnLeuValPro 60
 Db 871 GGCGTCGCCCGGTCGCCCGCGCGCGCGCTGCCCGCGCGAGAGCTTCTTGTTCG 812
 Qy 61 GlnArgLeuLeuValGlnHisValAlaValLeuHisArgAspLeuAlaValGlnValTrp 80
 Db 811 CAGCGCTCTTGTGACGAGTCCGAGTACTGACCGAGAGCTTGGCGTGCAGTCTG 752
 Qy 81 AlaHisLeuAlaGlyGlnLeuGlyHisGlyGluGluArgAlaGluHisLeuValGlnAla 100
 Db 751 GCTCATCTCGCGAGCTTGGCGATGGCGAGAGCGCGGAGAACATCTGGTCCAGGCT 692
 Qy 101 AlaValLeuLeuGlyArgAspLeuGluValGlyAlaAlaLeuGlyValAlaHisGln 120
 Db 691 GCTGTCTTCTTGGCGGAGATCTCGAAGTAGCGCAGCGCTGGGGGTGCTGCCCCAG 632
 Qy 121 LeuLeuAspLeuAlaLeuValHisLeuAlaValGluValAlaValThrLeuValAlaAla 140
 Db 631 CTGCTCGATCTCGCGTGTCTCACTCGCGTAGAGTACCGGTCACTTGTTCGCGCA 572
 Qy 141 AspAspGlnGlyHisValHisValLeuLeuGlyPheValLeuGluAlaArgLeuGlyVal 160
 Db 571 GATGACACAGGGCACGTCACGTTCTCTTGTGTTTGTCTTGAGGCAAGACTTGGTGT 512
 Qy 161 GluAspLeuLeuProGluProLeuHisLeuLeuGluGlyValAlaValAlaValThrGlu 180
 Db 511 GAGGATCTGCTGCGCTGAGCCGCTGACCTCTCGAAGAGTTCGCGGTTCACAGCTGAA 452
 Qy 181 HisGlnAspGluAsnValSerCysGluAspGlyGluAlaProHisGlyGlyGluArgVal 200
 Db 451 CACCAGGATGAACGCTCTCTGTGAGGATGAGAGCGCGCGCATGCGGGAGACGGTG 392
 Qy 201 ValAlaGlyArgValGluAspValGluLeuValAspLeuAlaAlaAspGlyValGluLeu 220
 Db 391 GTTCCGCGACGTGTGAGGATGTGAGTGTGAGTACCTCGCGCGGATGGAGTAGAACTT 332
 Qy 221 AlaValGluValLeuAspGlyArgValGlyValLeuGluAlaAlaGlyGluGluAla 240
 Db 331 GCGGTGGAAGTCTCGATGGTAGGCTGTAGCGTCTCTGAGAGCGCGGTGAGGAGCG 272
 Qy 241 ArgHisAspGlyArgLeuAlaHisLeuGlyArgAlaGluAspAspHisAlaIleAlaVal 260
 Db 271 CGACAGATGCCGCTTGTGCCACCTTGACGAGCGGAGGATGACCATGCGATGACGATT 212
 Qy 261 LeuGlyArgAspThrGlnLeuArgValAlaArgAlaHisLeuLeuAspHisArgGlyGln 280
 Db 211 CTTGCCGCGGATCTCAGTCCGAGTCTCGGCGACATCTTCTTGATCATCGCGGCGAG 152
 Qy 281 PheHis 282
 Db 151 TTTTCAT 146
 RESULT 4
 LOCUS BC018041/c
 DEFINITION Homo sapiens, RAS, dexamethasone-induced 1, clone MGC:26290
 IMAGE:479482, mRNA, complete cdb.
 ACCESSION BC018041
 VERSION BC018041.1 GI:17390075
 KEYWORDS MGC.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 1758)
 AUTHORS Strausberg,R.
 TITLE Direct Submission
 JOURNAL Submitted (03-DEC-2001) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 REMARK COMMENT
 Contact: MGC help desk
 Email: cgaps-re@mail.nih.gov
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki Toshiyuki and Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www-sbgc.stanford.edu>
 Contact: (Dickson, Mark) mc@paxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IPAK Plate: 32 Row: f Column: 11.

FEATURES

source

Location/Qualifiers

1..1758
 /organism="Homo sapiens"
 /db_xref="LocusID:51655"
 /db_xref="taxon:9606"
 /clone="MGC:26290 IMAGE:479482"
 /tissue_type="Brain, hypothalamus"
 /clone_lib="NIH_MGC_96"
 /lab_host="DH10B"
 /note="Vector: pBluescript"

CDS

213..1058
 /codon_start=1
 /product="RAS, dexamethasone-induced 1"
 /protein_id="AAH18041.1"
 /db_xref="GI:17390076"
 /translation="MKLAAMTKMCPDSELSILPAKNCYRMVILGSSKVGKTAIVSRP
 LTGFEDATYPTIEDFKFYSIRGEVQLDILDTSGNHPFPAMRRLSILTDVFIIV
 FSLNDRDSFEVQRLQIILDKSLANKTKENVDPVLVLCGNKGRDRDFVREVDQREI
 EQLVDDPFCAYFEISAKNSSLQDMFRALFAMAKLPSEMSPLHRKVSQYCDVLH
 KKALRNKLLRAGSGGGGDPGDFAGIVAFAPRRPVSVHSDUMYIREKASAGSQADKE
 RCVIS"

BASE COUNT 387 a 576 c 501 g 294 t

ORIGIN

Alignment Scores:
 Pred. No.: 2,33e-58 Length: 1758
 Score: 1411.00 Matches: 282
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0

US-09-709-103-1P4 (1-282) x BC018041 (1-1758)

Qy 1 LeuAlaAspAlaAlaLeuLeuValLeuGlyLeuAlaAlaGlyAlaGlyLeuLeuAla 20
 Db 1058 CTAGCTGATGACGACGCGCTCTTGTCTTGGCTTGGCTGCGCGCTTCTCGG 999
 Qy 21 AspValHisGluValAlaValTyrAlaGlyProAlaArgGluGlyCysHisAspAlaLys 40
 Db 998 GATGTACATGAGTTCGCTGTACGCTGGCGCGCGCGCGAGAGGTGCCACGATGCCAA 939
 Qy 41 GlyValAlaArgValAlaAlaAlaAlaAlaAlaAlaGlyProGlnGlnLeuValPro 60
 Db 938 GGCGTCGCCCGGTCGCCCGCGCGCGCTGCCCGCGCGAGAGCTTCTTGTTCG 879
 Qy 61 GlnArgLeuLeuValGlnHisValAlaValLeuHisArgAspLeuAlaValGlnValTrp 80
 Db 878 CAGCGCTCTTGTGACGACGCTGCGAGTACTGACCGAGAGCTTGGCGGTGAGTCTG 819
 Qy 81 AlaHisLeuAlaGlyGlnLeuGlyHisGlyGluGluArgAlaGluHisLeuValGlnAla 100
 Db 818 GCTCATCTCGCTGGGCGAGCTTGGCCATGGCGAGAGCGCGCGAGCATCTGGTCCAGGCT 759

Contact: sapiens@watson.wustl.edu
 ----- Summary Statistics
 Center Project name: H_NH0524F11

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RPL11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P.Y., Zhao, B., Frenken, E., Tateno, M., Catanese, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.rsgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)

VECTOR: pBACE3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is AC073621. Actual start of this clone is at base position 1 of RPL1-524F11; actual end is at base position 183334 of RPL1-524F11.

The sequence from position 1987 to 2253 was derived from PCR

product of RPL1-524F11 BAC DNA.

Location/Qualifiers

source

1. 183334
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="17"
 /map="17"
 /clone="RPL1-524F11"
 /clone_lib="RPC1-11"

misc_feature 1. .339
 /note="similar to EST AI672824 (NID:g4852555) w73e06.x1"
 misc_feature 339. .887
 /note="similar to EST AI968384 (NID:g5765202) wn02e05.x1"
 misc_feature 339. .738
 /note="similar to EST AI289781 (NID:g3931622) qw12a07.x1"
 misc_feature 339. .658
 /note="similar to EST AW072137 (NID:g6027135) wz98e07.x1"
 misc_feature 340. .767
 /note="similar to EST AI014398 (NID:g3228779) ou93b12.s1"
 misc_feature 340. .686
 /note="similar to EST AW237782 (NID:g6570171) xm81e08.x1"
 misc_feature 346. .742
 /note="similar to EST AI359064 (NID:g4110685) qy25e07.x1"
 misc_feature 405. .884
 /note="similar to EST AI279211 (NID:g3917445) qm24c09.x1"
 misc_feature 782. .1335
 /note="similar to EST BF042464 (NID:g10759519)"
 misc_feature 822. .1319
 /note="similar to EST BE864687 (NID:g10385980)"
 misc_feature 834. .2931
 /note="CpG island (GC=69.0, o/e=0.80, #CpGs=226)"
 misc_feature 1135. .1585
 /note="similar to EST BF555582 (NID:g11665312)"

misc_feature 1168. .1583
 /note="similar to EST AW028127 (NID:g5886883) ww26c07.x1"
 misc_feature 1432. .1583
 /note="similar to EST BE390509 (NID:g9335874)"
 misc_feature 1440. .1583
 /note="similar to EST AW915326 (NID:g8081017)"
 misc_feature 1450. .1583
 /note="similar to EST AA790463 (NID:g2850583) vw04g12.r1"
 misc_feature 1795. .2255
 /note="similar to EST BE390509 (NID:g9335874)"
 misc_feature 1795. .2201
 /note="similar to EST AA790463 (NID:g2850583) vw04g12.r1"
 misc_feature 1795. .2153
 /note="similar to EST AW915326 (NID:g8081017)"
 misc_feature 1795. .2114
 /note="similar to EST AW028127 (NID:g5886883) ww26c07.x1"
 misc_feature 1804. .2071
 /note="similar to EST BE836084 (NID:g10268462)"
 misc_feature 1817. .2201
 /note="similar to EST AA672515 (NID:g2644732) vo59e02.r1"
 misc_feature 1844. .2201
 /note="similar to EST AW463433 (NID:g7033601)"
 misc_feature 1985. .2285
 /note="similar to EST AA852097 (NID:g2940690)"
 misc_feature 2538. .3033
 /note="similar to EST BF057610 (NID:g10811506)"
 misc_feature 2554. .2806
 /note="similar to EST BF195653 (NID:g11082770)"
 misc_feature 3561. .3924
 /note="similar to EST R18261 (NID:g771871) yg01a05.r1"
 repeat_region 4027. .4085
 /rpt_family="MER1_type"
 misc_feature 4670. .5117
 /note="similar to EST R42241 (NID:g817108) yf98b09.s1"
 misc_feature 4857. .5117
 /note="similar to EST T16275 (NID:g518437)"
 repeat_region 6155. .6307
 /rpt_family="Alu"
 repeat_region 6309. .6446
 /rpt_family="L2"
 repeat_region 6583. .6891
 /rpt_family="Alu"
 repeat_region 6960. .7153
 /rpt_family="L1"
 repeat_region 7525. .7838
 /rpt_family="Alu"
 repeat_region 7851. .8487
 /rpt_family="L1"
 repeat_region 9000. .9104
 /rpt_family="L2"
 misc_feature 11463. .11750
 /note="similar to EST BF344873 (NID:g11292331)"
 misc_feature 11465. .11750
 /note="similar to EST AI340185 (NID:g4077112) qo79d07.x1"
 misc_feature 11455. .11750
 /note="similar to EST AI571390 (NID:g4534764) tn45e12.x1"
 misc_feature 11465. .11750
 /note="similar to EST AI701138 (NID:g4989038) we10d09.x1"
 misc_feature 11465. .11750
 /note="similar to EST AI803637 (NID:g5369109) tc43d01.x1"
 misc_feature 11455. .11750
 /note="similar to EST BE747895 (NID:g10161887)"
 misc_feature 11465. .11750
 /note="similar to EST BF203429 (NID:g11097015)"
 misc_feature 11465. .11697
 /note="similar to EST AI869572 (NID:g5543540) wl98d03.x1"
 misc_feature 11467. .11750
 /note="similar to EST AI817466 (NID:g5436545) wk38g01.x1"
 misc_feature 11467. .11619
 /note="similar to EST AA367144 (NID:g2019483)"
 misc_feature 11468. .11750
 /note="similar to EST BE218333 (NID:g8905651) hv37c09.x1"
 misc_feature 11468. .11750

Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, S., Severy, P., Sogniez, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, W.

TITLE

Submitted (03-MAR-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE

3 (bases 1 to 183598)

AUTHORS

Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barina, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Collangelo, M., Collins, S., Collamore, A., Cooke, A., Cooke, P., DeArelano, K., Dewar, K., Diaz, J. S., Dodge, S., Fero, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hages, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., LaRocque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G., MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, W.

TITLE

Submitted (22-MAR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

COMMENT

On Mar 22, 2002 this sequence version replaced gi:1645176. All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L11985
Center clone name: 524_F_11

* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 178215: contig of 178215 bp in length
* 178216 178315: gap of 100 bp
* 178316 183598: contig of 5283 bp in length.

FEATURES

source

1. 183598
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="17"
/chromosome="17"

/clone="RP11-524F11"

/clone.lib="RPC1-11 Human Male BAC"

40950 a 47626 c 49185 g 45565 t 272 others

BASE COUNT

ORIGIN

Alignment Scores:

Pred. No. : 5.69e-52 Length: 183598

Score: 1314.00 Matches: 280
Percent Similarity: 78.87% Conservative: 0
Best Local Similarity: 78.87% Mismatches: 2
Query Match: 93.13% Indels: 73
DB: 2 Gaps: 1

US-09-709-103-1f4 (1-282) x AC090608 (1-183598)

QY 1 LeuAlaAspAspAlaAlaLeuValLeuGlyLeuAlaAlaGlyAlaGlyLeuLeuAla 20
Db 182566 CTAGCTGATGACGAGCGCTCTTGTCTTGGCTTGGCTCGCGCTTCTCGCG 182507

QY 21 AspValHisGluValAlaValTyAlaGlyProAlaArgGluGlyCysHisAspAlaLys 40
Db 182506 GATGTACATGAGTTCGTGTGTACGCTGGCGCGCGGAGAGGTGCGACAGATGCCAA 182447

QY 41 GlyValAlaArgValAlaAlaAlaAlaAlaAlaAlaAlaGlyProGlnGlnLeuValPro 60
Db 182446 GCGCTCGCGCGGTTCGCGCGCGCGCTCGCGCTCGCGCGCGCGCGCTTCTTGTTCG 182387

QY 61 GlnArgLeuLeuValGlnHisValAlaValLeuHisArgAspLeuAlaValGlnValTrp 80
Db 182386 CAGCGCTTCTTGTGACGACGCTGCGAGTACTGACCGAGACCTTCGCGTGCAGGTCTGG 182327

QY 81 AlaHisLeuAlaGlyGlnLeuGlyHisGlyGluGluArgAlaGluHisLeuValGlnAla 100
Db 182326 GCTCATCTCGCTGGCGAGCTTGGCCATGGCGAGAGCGCGCGGAAACATCTGGTCCAGGCT 182267

QY 101 AlaValLeuLeuGlyArgAspLeuGluValGlyAlaAlaLeuGlyValValAlaHisGln 120
Db 182266 GCTGTCTTCTTTCGCGGAGATCTCGAAGTAGCGCGAGCTGCGGGTCTGCGCCACCA 182207

QY 121 LeuLeuAspLeuAlaLeuValHisLeuAlaValGluValAlaValThrLeuValAlaAla 140
Db 182206 CTGCTCGATCTCGCTGTGTTCACCTCGCGGTAGAGTCCGCTACCTTGTTCGCGCA 182147

QY 141 AspAspGlnGlyHisValHis--ValLeuLeuGlyPheValLeuGluAlaArgLeuGly 160
Db 182146 GATCACCGAGGCGACCTCCNACGTNTCTCTTGGTTTCTTCTTGGGCAAGACTTGGTG 182087

QY 160 alGluAspLeuProGluProLeuHisLeuLeuGlyValAlaValValGlnThrG 180
Db 182086 TCAGGATCTGCTGCTGAGCGCTGCACCTCTCGAAGAGTTCGCGTGTTCACAGACTG 182027

QY 180 luhisGlnAspGluAsnValser----- 187
Db 182026 AACACCAGGATGAACAGTCTCTCTAGAGGGGCGACAGAGCAGAAAGGAGGTGAG 181967

QY 187 ----- 187

Db 181966 GGTGGCGCGCCAGGGGACAAGTCGGGCTGACTTTGAGGCGCGCGCGCGCTAGCCTC 181907

QY 187 ----- 187

Db 181906 TCTAAGCGGAAGCGCACTACTTCGCGCGCGCGACCGCGCTCGCGGGCGGCC 181847

QY 188 -----CysGlu 189
Db 181846 TGGCCGAGGTTTCCCGCCCTTCCCTCCCGCACCTCCCGCGCGCGCGCTACCTGTGAG 181787

QY 190 AspGlyGluAlaProHisGlyGlyGluArgValValAlaGlyArgValGluAspValGlu 209
Db 181786 GATGGAGAGCGCGCATGCGCGGGAACGGGTGGTTGCGGACGCTGTCAGGATGTCGAG 181727

QY 210 LeuValAspLeuAlaAlaAspGlyValGluLeuAlaValGluValLeuAspGlyArgArg 229
Db 181726 CTGGTAGACTCGCGCGGATGAGTAGAATTCGCGTGGAAATCCTCGATGGTAGGGCT 181667

QY 230 ValGlyValLeuGluAlaAlaGlyGluAlaArgHisAspGlyArgLeuAlaHisLeu 249
Db 181666 GTAGGCGTCTCGAAGCGCGGTGAGGAAGCGCGACACGATGCGCGTCTTGTCCACCTT 181607

QY 250 GlyArgAlaGluAspHisAlaAlaValLeuGlyArgAspThrGlnLeuArgVal 269

Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome
 Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www-shgc.stanford.edu>
 Contact: (Dickson, Mark) mcd@paxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
 R. M.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAC Plate: 58 Row: 9 Column: 5
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 6677672.

FEATURES

source

```

1. .1612
  /organism="Mus musculus"
  /db_xref="taxon:10090"
  /map="FVB/N"
  /clone="MGC:36188 IMAGE:4989312"
  /tissue_type="Colon, normal. 5 month old male mouse."
  /clone_lib="NCI CGAP_Co24"
  /lab_host="PH108"
  /note="Vector: pCMV-SPORT6"
108. .950
  /codon_start=1
  /product="RAS, dexamethasone-induced 1"
  /protein_id="AAH34166.1"
  /db_xref="GI:21706875"
  /db_xref="LocusID:19416"
  /translation="MKLAAMIKMCPDSELSIPAKNRYRVIIGSSKVGKTAIVSRF
  LTRFEDATPTIEDFHRKFSIRGEVYOLDLDTSGNHPFPPMRRLSILTDGVFIIV
  FSLNDRDSFEVRLKQQLDPTKCLKNKTNVDVPLVIGNKGDRDFYREVQREI
  EQLVGDPPQRCAYFISAKNSLDDQMFALFAMAKLPSEMSPLHRKVSVOYCDVLH
  KKALRNKKLLRAGSGGGDHGDAFGILLAPFARRPSVHSDLMYIREKTSVGSQAKDER
  CVIS"

```

BASE COUNT 417 a 424 c 420 g 351 t
 ORIGIN

Alignment Scores:
 Pred. No.: 8,76e-37 Length: 1612
 Score: 949.50 Matches: 200
 Percent Similarity: 79.43% Conservative: 24
 Best Local Similarity: 70.92% Mismatches: 57
 Query Match: 67.29% Indels: 1
 DB: 10 Gaps: 1

US-09-709-103-1F4 (1-282) x BC034166 (1-1612)

```

QY 1 LeuAlaAspAlaAlaLeuValLeuGlyLeuAlaAlaGlyAlaGlyLeuAla 20
Db 950 CTAACGTGACACACGCGCTCTTCTCTAGCCTGGCTGCCGACACGTGTTTTCACG 891
QY 21 AspValHisGluValAlaValTyrAlaGlyProAlaArgGluGlyCysHisAspAlaLys 40
Db 890 AATGTACATGAGTGTGCTGTGACGCTGGTCTGCCACCAAGGCGCCAGATGCCA 831
QY 41 GlyValAlaArgValAlaAlaAlaAlaAlaAlaGlyProGlnGlnLeuValPro 60
Db 830 GGCATCGCC--GTGGTGGCCCGCCCTCCGCTGCCGACGACGAGCTCTTGTTCCT 774
QY 61 GlnArgLeuValGlnHisValAlaValLeuHisArgAspLeuAlaValGlnValTrp 80
Db 773 CAGACGCTTCTTGTCAGTACGTCGACGATCTGACATACCTTGGCGGTGCAAGTCGG 714
QY 81 AlaHisLeuAlaGlyGlnLeuGlyHisGlyGluGluArgAlaGluHisLeuValGlnAla 100
Db 713 GCTCATCTCGCTAGGACGCTTGGCCATGGCAAGAGCGCACGGAACATCTGTTCCAGCT 654
QY 101 AlaValLeuLeuGlyArgAspLeuGluValGlyAlaAlaLeuGlyValValAlaHisGln 120

```

```

Db 653 GCTGTTCTTCTTGGCTGAGATCTCGAAAGTAGGCACAAACGCTGAGGCTGTCACCCACCAG 594
QY 121 LeuLeuAspLeuAlaLeuValHisLeuAlaValGluValAlaValThrLeuValAlaAla 140
Db 593 CTGCTCAATCTCCCGTCTCTACTTCCCGTAGAAGTCCCGGTCCTCTTGTGTACCGCA 534
QY 141 AspAspGlnGlyHisValHisValLeuLeuGlyPheValLeuGluAlaArgLeuGlyVal 160
Db 533 AATGACACAGCGGACGCTCCACATCTCTTGTGTTTCTTCTTGAGACAGGACTTGGTGT 474
QY 161 GluAspLeuLeuProGluProLeuHisLeuLeuGluGlyValAlaValValGlnThrGlu 180
Db 473 TAGGATCTGCTGTTTGGAGCCTTTCACCTCTTCAATGAGTCCGGTGTCTAAGCTGAA 414
QY 181 HisGlnAspGluAsnValSerCysGluAspGlyClyAlaProHisGlyGlyGluArgVal 200
Db 413 CACCAAGATGAAACGCTCTCTGTGAGAGTAGAGAGCGCGCATGGCGGGAACCGATG 354
QY 201 ValAlaGlyArgValGluAspValGluLeuValAspLeuAlaAlaAspGlyValGluLeu 220
Db 353 ATTCCGGGATGTGTCAGTATGTCACACTGTTAGCTTCCGCGGATCGAGTAAACCTT 294
QY 221 AlaValGluValLeuAspGlyArgValGlyValLeuGluAlaAlaGlyGluGluAla 240
Db 293 TCGGTGGAAGTCTTCGATGTGAGGGGTGTAAAGCATCTCTGAGACATCTTCTTATCATCGCGCCAG 234
QY 241 ArgHisAspGlyArgLeuAlaHisLeuGlyArgAlaGluAspHisAlaLeuAlaVal 260
Db 233 CGACACAATGGCCGCTCTTCCCACTTTGGATGAGCGCGGATGACCATCTCTGAGCAGTT 174
QY 261 LeuGlyArgAspThrGlnLeuArgValAlaAlaArgAlaHisLeuLeuAspHisArgGlyGln 280
Db 173 CTTGGCGCGGATATCTCAGTTTCAAGTCTGCTTGGGCACATCTTCTTATCATCGCGCCAG 114
QY 281 PheHis 282
Db 113 TTTCAT 108

RESULT 14
AF009246/c
LOCUS
DEFINITION Mus musculus ras-related protein (DEXRAS1) mRNA, complete cds.
ACCESSION AF009246
VERSION AF009246.1 GI:2253712
KEYWORDS
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE
AUTHORS Kempainen,R.J. and Behrend,E.N.
TITLE Dexamethasone rapidly induces a novel ras superfamily
JOURNAL member-related gene in AtT-20 cells
MEDLINE J. Biol. Chem. 273 (6), 3129-3131 (1998)
PUBMED 98123070
9452419
REFERENCE
AUTHORS Kempainen,R.J. and Behrend,E.N.
TITLE Direct Submission
JOURNAL Submitted (18-JUN-1997) Physiology & Pharmacology, Auburn
University College of Veterinary Medicine, 213 Greene Hall, Auburn,
AL 36849, USA
FEATURES
source
1. .1623
  /organism="Mus musculus"
  /db_xref="taxon:10090"
  /cell_line="corticotrope tumor cell line AtT-20"
  /gene="DEXRAS1"
  /note="induced by dexamethasone"
  /codon_start=1

```

```

/product="ras-related protein"
/protein_id="AAC53538.1"
/db_xref="GI:2253713"
/translation="MKLAAMTKMCPSPDSELSIPAKNCYRMVILGSSKVGKTAIVSHF
LTGRADAVPTIEDFHRRFYSIRGEVYOLDILPDSGKHPAPRRISILGDFVILV
FSLNDRSDSEERQRLKOQILDTKSLCKTKENVDVLPYCKNGDDPFRYEVORRT
EQLVDDDDQRCAFEISAKKNSLDQMRALFMAAKIPSEMSPLHKRVSYQYCDVIL
KKAIRKKLRLRAGSGGSDHDAFGILAPFARRPSVHSDIMVIREKTSVGSQAKDKER
CVIS"

BASE COUNT      412 a      437 c      417 g      357 t
ORIGIN
Alignment Scores:      1.87e-36      Length:      1623
Score:      942.50      Matches:      199
Percent Similarity:      79.08%      Conservative:      24
Best Local Similarity:      70.57%      Mismatches:      58
Query Match:      66.80%      Indels:      1
DB:      10      Gaps:      1

US-09-709-103-1F4 (1-282) x AF009246 (1-1623)
QY      1 LeuAlAspAspAlaAlaLeuLeuValLeuAlaAlaGlyAlaGlyLeuLeuAla 20
DB      984 CTAACTGATGACACAGCGCTCTGCTGCTTACCTGGCTGGCAGACTGTTTTCACG 925
QY      21 AspValHisGlyValAlaValAlaGlyProAlaArgGlyGlyCysHisAspAlaLys 40
DB      924 AATGATCATGAGTGGCTGTCAGCGTGGGTCTGGCAGAGCAAGGGCGCCAAATCCAAA 865
QY      41 GlyValAlaArgValAlaAlaAlaAlaAlaAlaGlyProGlnGlnLeuLeuValPro 60
DB      864 GGCATCGCC---GTGGTCCCGCCCGCTCCGCGCCGACAGCAAGCTTCTTCTTCT 808
QY      61 GlnArgLeuLeuValGlnHisValAlaValLeuHisArgAspLeuAlaValGlnAlaTrp 80
DB      807 CAGAGCTTCTTGTGAGTACGTCGAGTGCACAGTACCTTGGCGGTGCAATCGCG 748
QY      81 AlaHisLeuAlaGlyGlnLeuGlyHisGlyGlnGlyArgAlaGlnHisLeuValGlnAla 100
DB      747 GCTCATCTCGTAGGAGCTTGGCCATGCAAGCGGACCGAACATCTGCTCCAACT 688
QY      101 AlaValAlaLeuGlyArgAspLeuGlyValAlaAlaLeuGlyValAlaAlaHisGln 120
DB      687 GGTGTTCTTCTTGGTGAATCTCGAAGTAGGCACAAACGCTGAGGGTGTCTACCCACAG 628
QY      121 LeuLeuAspLeuAlaLeuValHisLeuAlaValGlnValAlaValThrLeuValAlaAla 140
DB      627 CTGCTCAATCTCCCGCTGCTCTTACTTCCCGGTAGAAGTCCCGGTCCCTTGTACCGCA 568
QY      141 AspAspGlnGlyHisValHisValLeuGlyGlyPheValLeuGlyAlaArgLeuGlyVal 160
DB      567 AATGACACAGCGGACGTCACATCTCTTGGTTTGTCTTGAACAGCAATTTGTTGTC 508
QY      161 GlnAspLeuLeuProGlnProLeuHisLeuLeuGlyValAlaValAlaGlnThrGln 180
DB      507 TAGGATCTGCTCTTGGAGGCTTGGACCTTTCGATGAGTGGCGGTGTCTAAAGCTGAA 448
QY      181 HisGlnAspGlnAspValSerCysGlnAspGlyGlnAlaProHisGlyGlnLysGlyAl 200
DB      447 CACCGAAGTGAAGAAAGCTCTCTGTGAGTGAAGAGCGCCCGCATGGCGGGAAGGATG 388
QY      201 ValAlaGlyArgValGlnAspValGlnLeuValAspLeuAlaAlaAspGlyValGlnLeu 220
DB      387 ATTGGCGGAGTGTGTCAGATGTCACACTGTAAGCTTCCCGCGGATCGATTAACACTT 328
QY      221 AlaValGlnValLeuAspGlyArgArgValGlyValLeuGlnAlaAlaGlyGlnGlnAla 240
DB      327 TCGGTGGAGAGTCTCGATGTAGGGGTGTAAAGCATCTCGAAACGCGCCGTGAGGAACG 268
QY      241 ArgHisAspGlyArgLeuAlaHisLeuGlyArgAlaGlnAspHisAlaAlaLeuAlaVal 260
DB      267 CAAACAAATGGCCGCTTGGCCCACTTGGATGAGCCGAGGATCAACCATCTGTGACAGTT 208

```

```

QY      261 LeuGlyArgAspThrGlnLeuArgValAlaAlaArgAlaHisLeuLeuAspHisArgGlyGln 280
DB      207 CTGGCGCGGATACTACTGTTCAAGAGTGGTGGCAGCATCTTTGATGATCTCGCGGACG 148
QY      281 PheHis 282
DB      147 TTTCAT 142

RESULT 15
AL603710      179124 bp      DNA      linear      ROD 05-APR-2002
LOCUS      AL603710
DEFINITION      Mouse DNA sequence from clone Rp23-247B13 on chromosome 11,
complete sequence.
ACCESSION      AL603710
VERSION      AL603710.8      GI:20068514
KEYWORDS      HTG.
SOURCE      house mouse.
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
1 Almeida, J.
Direct Submission
Submitted (04-APR-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Apr 7, 2002 this sequence version replaced gi:18070899.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em: EMBL; Sw:
SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep
from the RPI-23 Mouse PAC Library
constructed by the group of Pieter de Jong.
For further details see http://www.chori.org/bacpac/home.htm
VECTOR: pBAC3.6.

```

FEATURES

source

```

1..179124
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="11"
/clone="Rp23-247B13"
/clone_1lb="RPI-23"

```

```

BASE COUNT      46162 a      43228 c      43736 g      45998 t
ORIGIN

```

Alignment Scores:

```

Pred. No.:      2.98e-33      Length:      179124
Score:      914.00      Matches:      200
Percent Similarity:      67.27%      Conservative:      24
Best Local Similarity:      60.06%      Mismatches:      57
Query Match:      64.78%      Indels:      52
DB:      10      Gaps:      2

```

```

US-09-709-103-1F4 (1-282) x AL603710 (1-179124)

```

```

QY      1 LeuAlAspAspAlaAlaLeuLeuValLeuAlaAlaGlyAlaGlyLeuLeuAla 20
DB      121430 CTAACTGATGACACAGCGCTCTTGTCTTAACTGCTGGCAGCACTGTTTTCACG 121489

```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 30, 2002, 14:39:35 / Search time 737 Seconds

(without alignments)
155,438 Million cell updates/sec

Title: US-09-709-103-141

Perfect score: 1450

Sequence: 1 MKLAMIRKMPDSELSIP.....REKASAGQANDKRCVIS. 282

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 363474 seqs, 203117208 residues

Total number of hits satisfying chosen parameters: 726948

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -DEV=tblp
-MODE=frame_plus_p2n.model -DEV=tblp
-Q=/cg2_1/USFTO.spool/TRANSUS09709103/runat.30122002.143926.20706/app.query.fasta.1.455
-DB=Published_Applications_NA -QMT=fastcap -SUFFIX=p2n.rmp -MINMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=dlousum62
-TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR SCORE=spct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000
-USER=TRANSUS09709103 @CGN 1.1.36 @runat.30122002.143926.20706 -NCPU=6 -ICPU=3
-NO_XLPEXY -NO_WMAP -LARGEOUTERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Published_Applications_NA.*

1: /cg2_6/ptodata/1/pubna/US07_PUBCOMB.seq.*
2: /cg2_6/ptodata/1/pubna/PCT_NEW_PUB.seq.*
3: /cg2_6/ptodata/1/pubna/US06_NEW_PUB.seq.*
4: /cg2_6/ptodata/1/pubna/US06_PUBCOMB.seq.*
5: /cg2_6/ptodata/1/pubna/US07_NEW_PUB.seq.*
6: /cg2_6/ptodata/1/pubna/PCTUS_PUBCOMB.seq.*
7: /cg2_6/ptodata/1/pubna/US08_NEW_PUB.seq.*
8: /cg2_6/ptodata/1/pubna/US08_PUBCOMB.seq.*
9: /cg2_6/ptodata/1/pubna/US09_NEW_PUB.seq.*
10: /cg2_6/ptodata/1/pubna/US09_PUBCOMB.seq.*
11: /cg2_6/ptodata/1/pubna/US10_NEW_PUB.seq.*
12: /cg2_6/ptodata/1/pubna/US10_PUBCOMB.seq.*
13: /cg2_6/ptodata/1/pubna/US60_NEW_PUB.seq.*
14: /cg2_6/ptodata/1/pubna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Match Length	DB ID	Description
1	883 60.9	3082 10	US-09-778-963A-1	Sequence 1, Appl1
2	526 36.3	11221 10	US-09-778-963A-3	Sequence 3, Appl1
3	485 33.4	405 10	US-09-960-352-10273	Sequence 10273, A

4	370	25.5	368	10	US-09-864-761-21643	Sequence 21643, A
5	300	20.7	3346	9	US-09-764-868-67	Sequence 67, Appl
6	291	20.1	688	9	US-09-764-868-490	Sequence 490, Appl
7	283.5	19.6	2040	12	US-10-044-090-336	Sequence 336, App
8	278.5	19.2	197997	10	US-09-822-246-3	Sequence 3, Appl1
9	276.5	19.1	570	12	US-10-104-484-1	Sequence 1, Appl1
10	276.5	19.1	570	12	US-10-104-484-3	Sequence 25, Appl1
11	275.5	19.0	551	10	US-09-765-298A-25	Sequence 25, Appl1
12	272.5	18.8	570	10	US-09-765-298A-27	Sequence 27, Appl1
13	264.5	18.2	930	10	US-09-801-368-285	Sequence 285, App
14	259.5	17.9	536	10	US-09-867-701-5566	Sequence 5566, Ap
15	255.5	17.6	1271	12	US-10-044-090-110	Sequence 110, App
16	254	17.5	942	10	US-09-962-832-256	Sequence 256, App
17	249	17.2	847	10	US-09-822-849A-40	Sequence 40, Appl
18	245.5	16.9	969	10	US-09-801-368-287	Sequence 287, App
19	244	16.8	952	10	US-09-917-800A-1389	Sequence 1389, Ap
20	237	16.3	2418	9	US-09-764-868-84	Sequence 84, Appl1
21	237	16.3	2479	9	US-09-764-868-69	Sequence 69, Appl1
22	236.5	16.3	1085	9	US-10-067-813-3	Sequence 3, Appl1
23	236	16.3	1119	9	US-10-067-813-1	Sequence 1, Appl1
24	231.5	16.0	452	10	US-09-960-352-4253	Sequence 4253, Ap
25	226	15.6	1088	10	US-09-917-800A-1616	Sequence 1616, Ap
26	225.5	15.6	4454	10	US-09-960-352-5830	Sequence 5830, Ap
27	225	15.5	1191	12	US-10-044-090-111	Sequence 111, App
28	214.5	14.8	1296	9	US-09-954-531-983	Sequence 983, App
29	214.5	14.8	1296	9	US-09-954-531-1378	Sequence 1378, Ap
30	208	14.3	412	10	US-09-960-352-6794	Sequence 6794, Ap
31	206	14.2	1199	10	US-09-794-257-10	Sequence 10, Appl
32	203	14.0	1224	10	US-09-972-529-1	Sequence 1, Appl1
33	203	14.0	3192	10	US-09-788-654A-1	Sequence 1, Appl1
34	200.5	13.8	612	10	US-09-972-529-3	Sequence 3, Appl1
35	194.5	13.4	836	10	US-09-770-445-674	Sequence 674, App
36	194	13.4	1274	10	US-09-925-102-91	Sequence 91, Appl
37	193	13.3	1572	9	US-09-764-868-501	Sequence 501, Appl
38	193	13.3	1157	9	US-10-108-605-44	Sequence 44, Appl
39	192	13.2	401	10	US-09-960-352-6793	Sequence 6793, Ap
40	192	13.2	552	10	US-09-794-257-12	Sequence 12, Appl
41	190.5	13.1	884	10	US-09-770-445-515	Sequence 515, App
42	190	13.1	774	10	US-09-919-580-42	Sequence 42, Appl
43	187	12.9	666	9	US-09-938-842A-40	Sequence 420, App
44	185.5	12.8	2734	9	US-09-764-868-489	Sequence 489, App
45	185.5	12.8	2964	9	US-10-091-613-2	Sequence 2, Appl1

ALIGNMENTS

RESULT 1
US-09-778-963A-1
; Sequence 1, Application US/09778963A
; Parent No. US20020115172A1
; GENERAL INFORMATION:
; APPLICANT: NEELAM, Beena et al
; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE
; FILE REFERENCE: CL001112
; CURRENT APPLICATION NUMBER: US/09/778, 963A
; CURRENT FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3082
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-778-963A-1

Alignment Scores:

Pred. No.:	6,21e-102	Length:	3082
Score:	883.00	Matches:	173
Percent Similarity:	77.78%	Conservative:	44
Best Local Similarity:	62.01%	Mismatches:	46
Query Match:	60.90%	Indels:	16
DB:	10	Gaps:	4


```

US-09-709-103-1F1 (1-282) x US-09-778-963A-1 (1-3082)
QY 5 AlaMetIleLysLysMetCysProSerAspSerGluLeuSerIleProAlaLysAsnCys 24
DB 208 GCCATGATGAAGACTTTCAGCGGGAACTGACGCTCAGTGTGCCCGCCAAAATCTCA 267
QY 25 TyrArgMetValIleLeuGlySerSerLysValGlyLysThrAlaIleValSerArgPhe 44
DB 268 TACCGCATGGTGGTCTCGGTGGCTCTCGGTGGCAAGACTCCATCGTGTCTCGCTTC 327
QY 45 LeuThrGlyArgPheGluAspAlaTyrThrProThrIleGluAspPheHisArgLysPhe 64
DB 328 CTCATGCGCCGCTTGTAGACACAGTACACACCCACCATCGAGGACTTCCACCGTAAAGTA 387
QY 65 TyrSerIleArgGlyGluValTyrGlnLeuAspIleLeuAspThrSerGlyAsnHisPro 84
DB 388 TACAACATCCCGCGGACATGTACAGCTCGACATCTCTGGATACCTCTGGCAACACCC 447
QY 85 PheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIleLeuValPheSer 104
DB 448 TTCCCGCCATCGCGAGCTGTCCATCTCACAGGGATGTCTTCATCTCGTGGTGTTCAGC 507
QY 105 LeuAspAsnArgAspSerPheGluValGlnArgLeuArgGlnGlnIleLeuAspThr 124
DB 508 CTGGNATACCGGGAGTCTTCATGAGGTCAAGCGCTTCAGAGCAGATCTTGAGGTC 567
QY 125 LysSerCysLeuLysAsnLysThrLysGluAsnValAspValProLeuValIleCysGly 144
DB 568 AAGTCTCTGCTGAAGAACACAGAACAGAGCGCGGAGTGTCCATGGTTCATCTGTGGC 627
QY 145 AsnLysGlyAspArg--AspPheTyrArgGluValAspGlnArgGluIleGluLeu 163
DB 628 AACAAAGAACGACACCGGAGCTGTGCGCGCAGGTGCCACCCAGCGCGGAGCTGCTG 687
QY 164 ValGlyAspAspProGlnArgCysAlaTyrPheGluIleSerAlaLysLysAsnSerSer 183
DB 688 GTGTGCGGCGAC---GAGAACTCCGCTACTTCAGGTGTGCGGCAAGAGAACACCCAC 744
QY 184 LeuAspGlnMetPheArgAlaLeuPheAlaMetAlaLysLeuProSerGluMetSerPro 203
DB 745 GTGGACGAGATGTTCTACGTCTCTCAGCATGCGCAAGCTGCCACAGAGATGAGCC 804
QY 204 AspLeuHisArgLysValSerValGlnTyrCysAspValLeuHisLysLys-----Ala 221
DB 805 GCCCTGATCCAGATCTCCGTGAGTACGCTGACGCTTCCACCCCGGCGGCTTCTGTC 864
QY 222 LeuArgAsnLysLysLeuLeuArgAlaGlySerGlyGlyGlyGlyAspProGlyAsp 241
DB 865 ATGCGCGCTCAAGGAGATG-----GAC 888
QY 242 AlaPheGlyIleValAlaProPheAlaArgProSerValHisSerAspLeuMetTyr 261
DB 889 GCCTATGGCATGTTCTCGCCCTTCGCGCGCGCCAGCGTCAACAGTACCTCAAGTAC 948
QY 262 IleArgGluLysAlaSerAlaGlySerGlnAlaLysAspLysGluArgCysValIle 280
DB 949 ATCAAGGCCAAGGTCCTTCGGGAAGCGCGCGCTGAGAGGGCAAGTGCACCATC 1005

RESULT 2
US-09-778-963A-3
; Sequence 3, Application US/09778963A
; Patent No. US20020115172A1
; GENERAL INFORMATION:
; APPLICANT: NEELAM, Beena et al
; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
; FILE REFERENCE: CLO01112
; CURRENT APPLICATION NUMBER: US/09/778,963A
; CURRENT FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Fast-Seq for Windows Version 4.0

```

```

; SEQ ID NO 3
; LENGTH: 11221
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-778-963A-3
Alignment Scores:
Pred. No.: 9,36e-56 Length: 11221
Score: 526.00 Matches: 106
Percent Similarity: 72.40% Conservative: 33
Best Local Similarity: 55.21% Mismatches: 37
Query Match: 36.28% Indels: 16
DB: 10 Gaps: 4

US-09-709-103-1F1 (1-282) x US-09-778-963A-3 (1-11221)
QY 92 SerIleLeuThrGlyAspValPheIleLeuValPheSerLeuAspAsnArgAspSerPhe 111
DB 7681 TCTCTCCCTGCAGGGATGTCTTCATCTCTGGTGTTCACCTGGATACCGGAGTCTTTC 7740
QY 112 GluGluValGlnArgLeuArgGlnGlnIleLeuAspThrLysSerCysLeuLysAsnLys 131
DB 7741 GATGAGGTCAAGCGCTTCAGAAAGCAGATCTCTGGAGGTCAAGTCTCTGCTGAAGAACAG 7800
QY 132 ThrLysGluAsnValAspValProLeuValIleCysGlyAsnLysGlyAspArg---Asp 150
DB 7801 ACCAAGGAGCGCGGAGCTGCCATGTCTGTGTGCAACAAAGAACACCGCGCGAG 7860
QY 151 PheTyrArgGluValAspGlnArgGluIleGluGlnLeuValGlyAspAspProGlnArg 170
DB 7861 CTGTCCCGCCAGTGTCCACCCAGCGCGGAGTGTGTGTGCGGCGAC---GAGAAC 7917
QY 171 CysAlaTyrPheGluIleSerAlaLysLysAsnSerSerLeuAspGlnMetPheArgAla 190
DB 7918 TCGGCTACTCTGAGGTGTGCGCCAAAGAAACACCAACGTGGACGAGATGTTCTACGTG 7977
QY 191 LeuPheAlaMetAlaLysLeuProSerGluMetSerProAspLeuHisArgLysValSer 210
DB 7978 CTCTTCAGCATGGCCAAAGTGTCCACACGAGATGAGCCCGCCCTGTCATCGAAGATCTCC 8037
QY 211 ValGlnTyrCysAspValLeuHisLysLys-----AlaLeuArgAsnLysLysLeuLeu 228
DB 8038 GTGCAAGTACGTGAGCGCTTCCACCCAGCGCTTCTGCATGCGCGCGCTCAAGAGATG 8097
QY 229 ArgAlaGlySerGlyGlyGlyGlyGlyAspProGlyAspAlaPheGlyIleValAlaPro 248
DB 8098 -----GACGCTATGCGATGTTCTCGGCC 8121
QY 249 PheAlaArgArgProSerValHisSerAspLeuMetTyrIleArgGluLysAlaSerAla 268
DB 8122 TTCGCGCGCGCGCGCGCTCAACAGTCAAGTCAAGTACATCAAGGCCAAGGTCCTTTCGG 8181
QY 269 GlySerGlnAlaLysAspLysGluArgCysValIle 280
DB 8182 GAAGCCAGCGCGCGTGGAGAGGCAAGTGCACCATC 8217

RESULT 3
US-09-960-352-10273
; Sequence 10273, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 16511.008/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 10273
; LENGTH: 405

```

```

TYPE: DNA
ORGANISM: Bos taurus
FEATURE:
NAME/KEY: unsure
LOCATION: (3), (23), (31)
OTHER INFORMATION: unsure at all n locations
OTHER INFORMATION: Clone ID: 44-LHB3058-050-Q1-K1-C8
US-09-960-352-10273

Alignment Scores:
Pred. No.: 9,15e-53 Length: 405
Score: 485.00 Matches: 95
Percent Similarity: 89.19% Conservative: 4
Best Local Similarity: 85.59% Mismatches: 12
Query Match: 33.45% Indels: 0
DB: 10 Gaps: 0

US-09-709-103-1f1 (1-282) x US-09-960-352-10273 (1-405)
QY 1 MetLysLeuAlaIaMetIleLysMetCysProSerAspSerGluLeuSerIlePro 20
DB 72 ATGAATCTGCGCGATGATCAAGAAATGTGCCGACGACCTCGATGTGAGATCCCG 131
QY 21 AAlaYasnCyStyIraMetValIleLeuGlySerSerIleValGlySerThrAlaIle 40
DB 132 GCCAAGAACTGCTACCCGATGAGTGTCTCTGCTCTCCAGAGGTGGCAAGACGGCCATC 191
QY 41 ValSerArpPheLeuThrGlyArqPheGluAspAlaIleThrProThrIleGluAspPhe 60
DB 192 GTGTGGGCTTCTCGACGGCGCCGTACGAGACGCCCTACGCCCACTCGAAGACTTC 251
QY 61 HisArgLysPheIleYserIleArgGlyGlyValIleGlnLeuAspIleLeuAspThrSer 80
DB 252 CACCGCAAGTCTCTACGATCGATCGAGGAGGATCCAGTCACTCTCGACACGATAC 311
QY 81 GlysHisProPheProAlaMetArqIleuSerIleLeuThrGlyAspValPheIle 100
DB 312 GGCAACCGCGCATTCGCCCATCGAGCGCTCTGCTGACTTCTTACCGAGACGTTCTCATC 371
QY 101 LeuValPheSerLeuAspAsnArqAspSerPhe 111
DB 372 CTAGGGTTTCAGACAGTACACCGGACGCTTC 404

RESULT 4
US-09-864-761-21643
Sequence 21643, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
FILE REFERENCE: Aecmica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263,6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30

```

```

PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
SEQ ID NO 21643
LENGTH: 368
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AL022334.1
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.8
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.2
OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL = 3.3
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.6
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.9
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.9
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.7
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3
OTHER INFORMATION: NT HIT: g11418034, EVALUE 0.00e+00
OTHER INFORMATION: SWISSPROT HIT: P13856, EVALUE 8.00e-12
OTHER INFORMATION: EST_HUMAN HIT: BE389944.1, EVALUE 0.00e+00
US-09-864-761-21643

Alignment Scores:
Pred. NO.: 3,13e-38 Length: 368
Score: 370.00 Matches: 69
Percent Similarity: 89.01% Conservative: 12
Best Local Similarity: 75.82% Mismatches: 10
Query Match: 25.52% Indels: 0
DB: 10 Gaps: 0

US-09-709-103-1f1 (1-282) x US-09-864-761-21643 (1-368)
QY 5 AlaMetIleLysLysMetCysProSerAspSerGluLeuSerIleProAlaLysAsnCys 24
DB 95 GCCAATGGAAGACTTCTTCACGGGAGACTCGCATGAGTGTGGCCCGCAAAAATCA 154
QY 25 TyrArgMetValIleLeuGlySerSerIleValGlySerThrAlaIleValSerArgPhe 44
DB 155 TACCGCATGAGTGTCTGCTGCTCGGCTCGGCTGGGCAAGACTCCACGCTGCTCGCTTC 214
QY 45 LeuThrGlyArqPheGluAspAlaIleThrProThrIleGluAspPheHisArgLysPhe 64
DB 215 CTCATATGCGCGCTTGAAGACCAAGTACACCACTCCAGAGACTTCCACCGTAAGTA 274
QY 65 TyrSerIleArgGlyGlyValIleValIleGlnLeuAspIleLeuAspThrSerGlyAsnHisPro 84
DB 275 TACAAATCCGGCGGAGATGATGACAGCTCCAGATCTCGATCTGCGAACCACCC 334
QY 85 PheProAlaMetArqArgLeuSerIleLeuThr 95
DB 335 TTCGCCCATGCGGAGGCTGTCTCATCTCTCA 367

```

RESULT 5

US-09-764-868-67
 ; Sequence 67, Application US/09764868
 ; Patent No. US20020168711A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ; FILE REFERENCE: PT232
 ; CURRENT APPLICATION NUMBER: US/09/764,868
 ; CURRENT FILING DATE: 2001-01-17
 ; Prior application data removed - refer to PALM or file wrapper
 ; NUMBER OF SEQ ID NOS: 1510
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 67
 ; LENGTH: 3346
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: SITE
 ; LOCATION: (2787)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; NAME/KEY: SITE
 ; LOCATION: (2795)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; US-09-764-868-67

Alignment Scores:
 Pred. No.: 7,13e-28 Length: 3346
 Score: 300.00 Matches: 68
 Percent Similarity: 59.28% Conservative: 31
 Best Local Similarity: 40.72% Mismatches: 52
 Query Match: 20.69% Indels: 16
 DB: 9 Gaps: 3

US-09-709-103-1f1 (1-282) x US-09-764-868-67 (1-3346)

QY 25 TyrArgMetValIleLeuGlySerLysValGlyLysThrAlaIleValSerArgPhe 44
 Db 201 TACAAGGTAGTGTGGTGGAGTGGAGGGTTGGCAATCTGCCCTTACTGTGCAGTTT 260
 QY 45 LeuThrGlyArgPheGluAspAlaThrProThrIleGluAspPheHisArgLysPhe 64
 Db 261 GTCACTGGGACTTTCATTGAGNAATATGACCCCACTTGAAGATTCTACCGCAAGAG 320
 QY 65 TyrSerIleArgGlyGluValTyrGlnLeuAspIleLeuAspThrSerGlyAsnHisPro 84
 Db 321 ATCGAAGTGGACTCTTCCCTCCCTGCTGCTGGAATTTCTGGACACCGCAGGACTGAGCAG 380
 QY 85 PheProAlaMetArgLeuSerIleLeuThrGlyAspValPheIleLeuValPheSer 104
 Db 381 TTTCCTCCATGAGAGATCTTACATCAAAACCGCCCAAGGTTTCATCCTGTGTTATAGC 440
 QY 105 LeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGlnGlnIleLeuAspThr 124
 Db 441 CTGGTTAATCAACAGTCTTTTCAGGATATCAAGCAATGAGAGATCAAAATGTGC----- 494
 QY 125 LysSerCysLeuLysAsnLysThrLysGluAsnValAspValProLeuValIleCysGly 144
 Db 495 -----AGAGTGAAGAGATATGAAAAAGTCCCACTAATCCTAGTAGGA 536
 QY 145 AsnLysGlyAspArgAspPheTyrArgGluVal-----AspLysGluGluGlu 161
 Db 537 AATAAAGTGGATCTGGACACAGAAAGAGAGGTTATGTCTTCAGAGGCGAGAGCTCTGGCT 596
 QY 162 GlnLeuValGlyAspAspProGlnArgCysAlaTyrPheGluIleSerAlaLysLysAsn 181
 Db 597 CAAGAATGGGCG-----TGTCCTTTCATGGAGACATCGGCAAAAAGTAA 641
 QY 182 SerSerLeuAspGlnMetPhe 188
 Db 642 TCAATGTGGATGAATCTTTT 662

RESULT 6

US-09-764-868-490
 ; Sequence 490, Application US/09764868
 ; Patent No. US20020168711A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ; FILE REFERENCE: PT232
 ; CURRENT APPLICATION NUMBER: US/09/764,868
 ; CURRENT FILING DATE: 2001-01-17
 ; Prior application data removed - refer to PALM or file wrapper
 ; NUMBER OF SEQ ID NOS: 1510
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 490
 ; LENGTH: 688
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: SITE
 ; LOCATION: (579)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; NAME/KEY: SITE
 ; LOCATION: (610)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; NAME/KEY: SITE
 ; LOCATION: (669)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; US-09-764-868-490

Alignment Scores:
 Pred. No.: 8,72e-28 Length: 688
 Score: 291.00 Matches: 67
 Percent Similarity: 57.93% Conservative: 28
 Best Local Similarity: 40.85% Mismatches: 59
 Query Match: 20.07% Indels: 10
 DB: 9 Gaps: 2

US-09-709-103-1f1 (1-282) x US-09-764-868-490 (1-688)

QY 25 TyrArgMetValIleLeuGlySerLysValGlyLysThrAlaIleValSerArgPhe 44
 Db 194 TACAAGGTAGTGTGGTGGAGTGGAGGGTTGGCAATCTGCCCTTACTGTGCAGTTT 253
 QY 45 LeuThrGlyArgPheGluAspAlaThrProThrIleGluAspPheHisArgLysPhe 64
 Db 254 GTCACTGGGACTTTCATTGAGNAATATGACCCCACTTGAAGATTCTACCGCAAGAG 313
 QY 65 TyrSerIleArgGlyGluValTyrGlnLeuAspIleLeuAspThrSerGlyAsnHisPro 84
 Db 314 ATCGAAGTGGACTCTTCCCTCCCTGCTGCTGGAATTTCTGGACACCGCAGGAACTGAGCAG 373
 QY 85 PheProAlaMetArgLeuSerIleLeuThrGlyAspValPheIleLeuValPheSer 104
 Db 374 TTTCCTCCATGAGAGATCTTACATCAAAACCGCCCAAGGTTTCATCCTGTGTTATAGC 433
 QY 105 LeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGlnGlnIleLeuAspThr 124
 Db 434 CTGGTTAATCAACAGTCTTTTCAGGATATCAAGCAATGAGAGATCAAAATGTGC----- 487
 QY 125 LysSerCysLeuLysAsnLysThrLysGluAsnValAspValProLeuValIleCysGly 144
 Db 488 -----AGAGTGAAGAGATATGAAAAAGTCCCACTAATCCKAGTAGGA 529
 QY 145 AsnLysGlyAspArgAspPheTyrArgGluValAspGlnArgGluIleGluGlnLeuVal 164
 Db 530 AATAAAGTGGATCTGGACACAGAAAGAGAGGTTATGTCTTCAGAGGCGAGAGCTCTGGCT 589
 QY 165 GlyAspAspProGlnArgCysAlaTyrPheGluIleSerAlaLysLysAsnSerSerLeu 184
 Db 590 CAAGAA-----TGGGGCTGTCTCTTATGAGACATCGGCAAAAAGTAAATCAATGCTG 643
 QY 185 AspGlnMetPhe 188
 Db 644 GATGAATCTTTT 655

RESULT 7
 US-10-044-090-336
 ; Sequence 336, Application US/10044090
 ; Patent No. US20020137081A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Olga Bandman
 ; TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
 ; FILE REFERENCE: PA-0028 US
 ; CURRENT APPLICATION NUMBER: US/10/044,090
 ; CURRENT FILING DATE: 2002-01-09
 ; NUMBER OF SEQ ID NOS: 850
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 336
 ; LENGTH: 2040
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; OTHER INFORMATION: Incyte ID No. US20020137081A1 3110662CB1
 US-10-044-090-336

Alignment Scores:
 Pred. No.: 4,14e-26 Length: 2040
 Score: 283.50 Matches: 66
 Percent Similarity: 55.21% Conservative: 40
 Best Local Similarity: 34.38% Mismatches: 69
 Query Match: 19.55% Indels: 17
 DB: 12 Gaps: 3

US-09-709-103-1f1 (1-282) x US-10-044-090-336 (1-2040)

OY 25 TyrArgMetValIleLeuGlySerSerLySValGlyLeuThrAlaIleValSerArgPhe 44
 Db 202 TATAAGCTAGTCTCTTGGCTCAGAGCGCTTGAAGCTCTTGAAGCTAGTCAATTT 261
 OY 45 LeuThrGlyArgPheGluAspAlaIleThrProThrIleGluAspPheHisArgGlyPhe 64
 Db 262 GTTCAAGGAATTTTGTGAAATAATGATCTTCAATGAAATTTCTTATGAAAGCAA 321
 OY 65 TyrSerIleArgGlyGluValTyrGlnLeuAspIleLeuAspThrSerGlyAsnHisPro 84
 Db 322 GTTGAAGTAGATGACACAGAGCTATGCTTGAATCTTGGATCTGACGAAACGAGCAA 381
 OY 85 PheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIleLeuValPheSer 104
 Db 382 TTTCACGCAATGAGGATTTATACATGAAATAATGACAAAGATTGCAATGATTATTC 441
 OY 105 LeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGlnIleLeuAspThr 124
 Db 442 ATCACACGACAGTCCACATTTAAACGATTACAGACCTGAGAAACAGATTCTT----- 495
 OY 125 LysSerCysLeuLysAsnLysThrLysGluAsnValAspValProLeuValIleCysGly 144
 Db 496 -----CGAGTTAAAGCACTGATGATGCTTCAATGATTTCTTGTGGT 537
 OY 145 AsnLysGlyAspArgAspPheTyrArgGluValAspGlnArgGluIleGluGlnLeuVal 164
 Db 538 AATTAATGTGACTTGGA-----GATGAAAGAGTTGTAGGGAAGAACAA 582
 OY 165 GlyAspAspProGlnArg-----CysAlaTyrPheGluIleSerAlaLysLys 180
 Db 583 GGTCAAAATCTTACAGAAAGAAATGAAACAACTGTGCAATCTTGAATCTTGGCAAAATCA 642
 OY 181 AsnSerSerLeuAspGlnMetPheArgAlaLeuPheAlaMetAlaLysLeuProSerGlu 200
 Db 643 AAAATTAATGTTAATGAGATCTTTATGACCTAGCGGCAATTAACGAAAAACTCCA 702
 OY 201 MetSerProAspLeuHisArgLysValSerValGln 212
 Db 703 GTGCTGGGAAGGCTCGCAAAAGATCATCATGTCTAG 738
 RESULT 8

US-09-822-246-3/c
 ; Sequence 3, Application US/09822246
 ; Patent No. US20020142383A1
 ; GENERAL INFORMATION:
 ; APPLICANT: MERKULOV, Gennady et al.
 ; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
 ; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
 ; FILE REFERENCE: C0001149
 ; CURRENT APPLICATION NUMBER: US/09/822,246
 ; CURRENT FILING DATE: 2001-04-02
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 3
 ; LENGTH: 197997
 ; TYPE: DNA
 ; ORGANISM: Human
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (1) (197997)
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-822-246-3

Alignment Scores:
 Pred. No.: 2.01e-22 Length: 197997
 Score: 278.50 Matches: 65
 Percent Similarity: 55.21% Conservative: 41
 Best Local Similarity: 33.85% Mismatches: 69
 Query Match: 19.21% Indels: 17
 DB: 10 Gaps: 3

US-09-709-103-1f1 (1-282) x US-09-822-246-3 (1-197997)

OY 25 TyrArgMetValIleLeuGlySerSerLySValGlyLeuThrAlaIleValSerArgPhe 44
 Db 43187 TATAAGCTAGTCTCTTGGCTCAGAGCGCTTGAAGCTCTTGAAGCTAGTCAATTT 43128
 OY 45 LeuThrGlyArgPheGluAspAlaIleThrProThrIleGluAspPheHisArgGlyPhe 64
 Db 43127 GTTCAAGGAATTTTGTGAAATAATGATCTTCAATGAAATTTCTTATGAAAGCAA 43068
 OY 65 TyrSerIleArgGlyGluValTyrGlnLeuAspIleLeuAspThrSerGlyAsnHisPro 84
 Db 43067 GTTGAAGTAGATGACACAGAGCTATGCTTGAATCTTGGATCTGACGAAACGAGCAA 43008
 OY 85 PheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIleLeuValPheSer 104
 Db 43007 TTTCACGCAATGAGGATTTATACATGAAATAATGACAAAGATTGCAATGATTATTC 42948
 OY 105 LeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGlnIleLeuAspThr 124
 Db 42947 ATCACACGACAGTCCACATTTAAACGATTACAGACCTGAGAAACAGATTCTT----- 42894
 OY 125 LysSerCysLeuLysAsnLysThrLysGluAsnValAspValProLeuValIleCysGly 144
 Db 42893 -----CGAGTTAAAGCACTGATGATGCTTCAATGATTTCTTGTGGT 42852
 OY 145 AsnLysGlyAspArgAspPheTyrArgGluValAspGlnArgGluIleGluGlnLeuVal 164
 Db 42851 AATTAATGTGACTTGGA-----GATGAAAGAGTTGTAGGGAAGAACAA 42807
 OY 165 GlyAspAspProGlnArg-----CysAlaTyrPheGluIleSerAlaLysLys 180
 Db 42806 GGTCAAAATCTTACAGAAAGAAATGAAACAACTGTGCAATCTTGAATCTTGGCAAAATCA 42747
 OY 181 AsnSerSerLeuAspGlnMetPheArgAlaLeuPheAlaMetAlaLysLeuProSerGlu 200
 Db 42746 AAAATTAATGTTAATGAGATCTTTATGACCTAGCGGCAATTAACGAAAAACTCCA 42687
 OY 201 MetSerProAspLeuHisArgLysValSerValGln 212
 Db 42686 GTGCTGGGAAGGCTCGCAAAAGATCATCATGTCTAG 42651

RESULT 9

US-10-104-484-1
; Sequence 1, Application US/10104484
; Patent No. US20020150566A1
; GENERAL INFORMATION:
; APPLICANT: Stewart, Scott
; APPLICANT: Guan, Kun-Lian
; TITLE OF INVENTION: METHOD OF INHIBITING CANCEROUS CELL PROLIFERATION USING RAS MUTANT
; TITLE OF INVENTION: GDP-BOUND CONFORMATION
; FILE REFERENCE: UMICH-0010
; CURRENT APPLICATION NUMBER: US/10/104,484
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER: 60/277,959
; PRIOR FILING DATE: 2001-03-23
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 1
; LENGTH: 570
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Nucleotide sequence of the GDP-bound Ras mutant, RasN17N69.
US-10-104-484-1

Alignment Scores:

Pred. No.: 4,53e-26 Length: 570
Score: 276.50 Matches: 72
Percent Similarity: 52.74% Conservative: 34
Best Local Similarity: 35.82% Mismatches: 73
Query Match: 19.07% Indels: 23
DB: 12 Gaps: 6

US-09-709-103-1f1 (1-282) x US-10-104-484-1 (1-570)

QY 25 TyrArgMetValIleLeuGlySerSerLysValGlyLysThrAlaIleValSerArgPhe 44
DB 10 TATAAGCTGGTGGTGGCGCGCGGTGTGGCAAGATCGCTGACCATCCAGCTG 69
QY 45 LeuThrGlyArgPheGluAspAlaTyrThrProThrIleGluAspPheHisArgLysPhe 64
DB 70 ATCCAGAACCATTTTGTGGAGCAATACACCCCTATAGAGATCTCTACCGGAGCAG 129
QY 65 TyrSerIleArgGlyGluValTyrGlnLeuAspIleLeuAspThrSerGlyAsnHisPro 84
DB 130 GTGGTCATTGATGGGAGACGCGCTGTGGACATCTCTGGATACCGCGGCGAGGAG 189
QY 85 PheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIleLeuValPheSer 104
DB 190 TACAGCGCCATCGGAACACGACGATGCGACCGCGGAGGCGCTTCTGTGTGTGTGCC 249
QY 105 LeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGlnGlnIleLeuAspThr 124
DB 250 ATCAACACACCAAGTCTTTTGGAGCATCCACAGTACAGGAGCAGATC----- 300
QY 125 LysSerCysLeuLysAsnLysThrLysGluAsnValAspValProLeuValIleCysGly 144
DB 301 -----AAACGGGTGAAGACTCGATGACGTCGCCATGCTGTGTGGG 345
QY 145 AsnLysGlyAspArgAspPheTyrArgGluValAspGlnArgGluIleGluGlnLeuVal 164
DB 346 AACAAAGTGTGACCTG---GCTGCACGCACTGTGGAACTCTCGCAGGCTCAGGACCTGCC 402
QY 165 GlyAspAspProGlnArgCysAlaTyrPheGluIleSerAlaLysLysAsnSerSerLeu 184
DB 403 CGAAGC-----TACGGCATCCCTACATCGAGACTCGGCAAGACCCGCGGAGGAGTG 456
QY 185 AspGlnMetPheArgAlaLeu-----PheAlaMetAlaLysLeu--- 197
DB 457 GAGGATGCTCTTACACGTGTGGTGGTGGTGGATCCGCGACACCAAGCTCGGAAGCTGAAC 516
QY 198 ---ProSerGluMetSerProAspLeuHisArgLysValSerValGlnTyrCysAspVal 216
DB 517 CCTCTCTGATGAGTGGCCCCGG-CTGCAT-----GAGCTGAAGTGTGTCTCTC 566

QY 217 Leu 217
DB 567 CTG 569

RESULT 10

US-10-104-484-3
; Sequence 3, Application US/10104484
; Patent No. US20020150566A1
; GENERAL INFORMATION:
; APPLICANT: Guan, Kun-Lian
; APPLICANT: Stewart, Scott
; TITLE OF INVENTION: METHOD OF INHIBITING CANCEROUS CELL PROLIFERATION USING RAS MUTANT
; TITLE OF INVENTION: GDP-BOUND CONFORMATION
; FILE REFERENCE: UMICH-0010
; CURRENT APPLICATION NUMBER: US/10/104,484
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER: 60/277,959
; PRIOR FILING DATE: 2001-03-23
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 3
; LENGTH: 570
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Nucleotide sequence of the wild type Ras mutant, RasN17N69.
US-10-104-484-3

Alignment Scores:

Pred. No.: 4,53e-26 Length: 570
Score: 276.50 Matches: 72
Percent Similarity: 52.74% Conservative: 34
Best Local Similarity: 35.82% Mismatches: 73
Query Match: 19.07% Indels: 23
DB: 12 Gaps: 6

US-09-709-103-1f1 (1-282) x US-10-104-484-3 (1-570)

QY 25 TyrArgMetValIleLeuGlySerSerLysValGlyLysThrAlaIleValSerArgPhe 44
DB 10 TATAAGCTGGTGGTGGCGCGCGGTGTGGCAAGATCGCTGACCATCCAGCTG 69
QY 45 LeuThrGlyArgPheGluAspAlaTyrThrProThrIleGluAspPheHisArgLysPhe 64
DB 70 ATCCAGAACCATTTTGTGGAGCAATACACCCCTATAGAGATCTCTACCGGAGCAG 129
QY 65 TyrSerIleArgGlyGluValTyrGlnLeuAspIleLeuAspThrSerGlyAsnHisPro 84
DB 130 GTGGTCATTGATGGGAGACGCGCTGTGGACATCTCTGGATACCGCGGCGAGGAG 189
QY 85 PheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIleLeuValPheSer 104
DB 190 TACAGCGCCATCGGAACACGACGATGCGACCGCGGAGGCGCTTCTGTGTGTGTGCC 249
QY 105 LeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGlnGlnIleLeuAspThr 124
DB 250 ATCAACACACCAAGTCTTTTGGAGCATCCACAGTACAGGAGCAGATC----- 300
QY 125 LysSerCysLeuLysAsnLysThrLysGluAsnValAspValProLeuValIleCysGly 144
DB 301 -----AAACGGGTGAAGACTCGATGACGTCGCCATGCTGTGTGGG 345
QY 145 AsnLysGlyAspArgAspPheTyrArgGluValAspGlnArgGluIleGluGlnLeuVal 164
DB 346 AACAAAGTGTGACCTG---GCTGCACGCACTGTGGAACTCTCGCAGGCTCAGGACCTGCC 402
QY 165 GlyAspAspProGlnArgCysAlaTyrPheGluIleSerAlaLysLysAsnSerSerLeu 184
DB 403 CGAAGC-----TACGGCATCCCTACATCGAGACTCGGCAAGACCCGCGGAGGAGTG 456
QY 185 AspGlnMetPheArgAlaLeu-----PheAlaMetAlaLysLeu--- 197

Db 457 GAGGATGCTTCTACACGTTGGTGGCTGAGATCCGGACAGACACAGCTCGGAAAGTGAAC 516
Qy 198 ---ProSerGluMetSerProAspLeuHisArgLysValSerValGlnTyrCysAspVal 216
Db 517 CCTCTGATGATGAGTGGCCCGG-CTGCAT-----GAGCTCAAGTGTGTCTCTC 566
Qy 217 Leu 217
Db 567 CTG 569
RESULT 11
US-09-765-298A-25
; Sequence 25, Application US/09765298A
; Patent No. US20020137017A1
; GENERAL INFORMATION:
; APPLICANT: ARONHEIM, AMI
; TITLE OF INVENTION: METHOD FOR DETECTION PROTEIN-PROTEIN INTERACTIONS AND A KIT THERE
; FILE REFERENCE: 108387.01
; CURRENT APPLICATION NUMBER: US/09/765,298A
; CURRENT FILING DATE: 2001-01-22
; PRIOR APPLICATION NUMBER: IL 125456
; PRIOR FILING DATE: 1998-07-22
; PRIOR APPLICATION NUMBER: IL 128017
; PRIOR FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 25
; LENGTH: 551
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-765-298A-25

Alignment Scores:
Pred. No.: 5,76e-26 Length: 551
Score: 275.50 Matches: 67
Percent Similarity: 53.48% Conservative: 33
Best Local Similarity: 35.83% Mismatches: 68
Query Match: 19.00% Indels: 19
DB: Gaps: 5
US-09-709-103-1f1 (1-282) x US-09-765-298A-25 (1-551)

Qy 25 TyrArgMetValIleLeuGlySerSerLysValGlyLysThrAlaIleValSerArgPhe 44
Db 4 TATAAAGCTGTGTGTGTGGCGCGCGCGGTGGCGCAAGAGCGCTGACCATCCAGCTG 63
Qy 45 LeuThrGlyArgPheGluAspAlaTyrThrProThrIleGluAspPheHisArgLysPhe 64
Db 64 ATCCAGAACCATTTTGTGACGAAATACGACCCCACTATAGAGGATTCCTACCGGAAGCAG 123
Qy 65 TyrSerIleArgGlyGluValTyrGlnLeuAspIleLeuAspThrSerGlyAsnHisPro 84
Db 124 GTGCTCATTTGATGGGGAACGTCCTGTGGACATCTCGATACCGCGGCGCAGAGAGAG 183
Qy 85 PheProAlaMetArgArgLysSerIleLeuThrGlyAspValPheIleLeuValPheSer 104
Db 184 TACAGCGGCATGCGGAGACAGTACATGCGCACCGGGGAGGCGCTTCTGTGTGTGGTCC 243
Qy 105 LeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGlnIleLeuAspThr 124
Db 244 ATCAACCAACCAACCACTTTTGGAGCATCCACGATCAGAGGAGCAATC----- 294
Qy 125 LysSerCysLeuLysAsnLysThrLysGluAsnValAspValProLeuValIleCysGly 144
Db 295 -----AAACGGGTGAAGAGCTCGATGACGTCGTCCTATGTCTGTGTGGG 339
Qy 145 AsnLysGlyAspArgAspPheTyrArgGluValAspGlnArgGluIleGlnIleLeuVal 164
Db 340 AACAAAGTGTGACTG---GCTGACGCACTGTGGAATCTCGGACAGGCTCAGAGACTCGCC 396
Qy 165 GlyAspAspProGlnArgCysAlaTyrPheGluIleSerAlaLysLysAsnSerSerLeu 184
Db 397 CGAAGC-----TACGGCATCCCTTACATCGAGACCTCGGCCAAGACCCGGCAGGAGTG 450

Qy 185 AspGlnMetPheArgAlaLeu-----PheAlaMetAlaLysLeu--- 197
Db 451 GAGGATGCTTCTACACGTTGGTGGCTGAGATCCGGACAGACAGCTCGGAAAGTGAAC 510
Qy 198 ---ProSerGluMetSerPro 203
Db 511 CCTCTGATGATGAGTGGCCCGG 531

RESULT 12
US-09-765-298A-27
; Sequence 27, Application US/09765298A
; Patent No. US20020137017A1
; GENERAL INFORMATION:
; APPLICANT: ARONHEIM, AMI
; TITLE OF INVENTION: METHOD FOR DETECTION PROTEIN-PROTEIN INTERACTIONS AND A KIT THERE
; FILE REFERENCE: 108387.01
; CURRENT APPLICATION NUMBER: US/09/765,298A
; CURRENT FILING DATE: 2001-01-22
; PRIOR APPLICATION NUMBER: IL 125456
; PRIOR FILING DATE: 1998-07-22
; PRIOR APPLICATION NUMBER: IL 128017
; PRIOR FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 27
; LENGTH: 570
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-765-298A-27

Alignment Scores:
Pred. No.: 1.46e-25 Length: 570
Score: 272.50 Matches: 72
Percent Similarity: 53.23% Conservative: 35
Best Local Similarity: 35.82% Mismatches: 72
Query Match: 18.79% Indels: 23
DB: Gaps: 6
US-09-709-103-1f1 (1-282) x US-09-765-298A-27 (1-570)

Qy 25 TyrArgMetValIleLeuGlySerSerLysValGlyLysThrAlaIleValSerArgPhe 44
Db 10 TATAAAGCTGTGTGTGTGGCGCGCGCGGTGGCGCAAGAGCGCTGACCATCCAGCTG 69
Qy 45 LeuThrGlyArgPheGluAspAlaTyrThrProThrIleGluAspPheHisArgLysPhe 64
Db 70 ATCCAGAACCATTTTGTGACGAAATACGACCCCACTATAGAGGATTCCTACCGGAAGCAG 129
Qy 65 TyrSerIleArgGlyGluValTyrGlnLeuAspIleLeuAspThrSerGlyAsnHisPro 84
Db 130 GTGCTCATTTGATGGGGAACGTCCTGTGGACATCTCGATACCGCGGCGCAGAGAGAG 189
Qy 85 PheProAlaMetArgArgLysSerIleLeuThrGlyAspValPheIleLeuValPheSer 104
Db 190 TACAGCGGCATGCGGAGACAGTACATGCGCACCGGGGAGGCGCTTCTGTGTGTGGTCC 249
Qy 105 LeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGlnIleLeuAspThr 124
Db 250 ATCAACCAACCAACCACTTTTGGAGCATCCACGATCAGAGGAGCAATC----- 300
Qy 125 LysSerCysLeuLysAsnLysThrLysGluAsnValAspValProLeuValIleCysGly 144
Db 301 -----AAACGGGTGAAGAGCTCGATGACGTCGTCCTATGTCTGTGTGGG 345
Qy 145 AsnLysGlyAspArgAspPheTyrArgGluValAspGlnArgGluIleGlnIleLeuVal 164
Db 346 AACAAAGTGTGACTG---GCTGACGCACTGTGGAATCTCGGACAGGCTCAGAGACTCGCC 402
Qy 165 GlyAspAspProGlnArgCysAlaTyrPheGluIleSerAlaLysLysAsnSerSerLeu 184
Db 403 CGAAGC-----TACGGCATCCCTTACATCGAGACCTCGGCCAAGACCCGGCAGGAGTG 456

```

QY 185 AspGlnMetPheArgAlaLeu-----PheAlaMetAlaLysLeu--- 197
Db 457 GAGAGCCTTCTACAGTTCGGTGGAGATCGGCGAGCACAGCTCGGGAAGCTGAAC 516
QY 198 ---ProSerGluMetSerProAspLeuHisArgLysValSerValGlnTyrCysAspVal 216
Db 517 CCTCTGATGAGAGTGCCCGCG-CTGCAT-----GAGCTGCAAGTGTGTGCTCTC 566

QY 217 Leu 217
Db 567 CTG 569

RESULT 13
US-09-801-368-285
; Sequence 285, Application US/09801368
; Patent No. US20020128250A1
; GENERAL INFORMATION:
; APPLICANT: Bueby, Robert
; APPLICANT: Cali, Brian
; APPLICANT: Hecht, Peter
; APPLICANT: Holtzman, Doug
; APPLICANT: Madden, Kevin
; APPLICANT: Maxon, Mary
; APPLICANT: Milne, Todd
; APPLICANT: No. US20020128250A1man, Thea
; APPLICANT: Royer, John
; APPLICANT: Salama, Sofie
; APPLICANT: Sherman, Amir
; APPLICANT: Silva, Jeff
; APPLICANT: Summers, Eric
; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
; FILE REFERENCE: 109272.147
; CURRENT APPLICATION NUMBER: US/09/801,368
; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 09/487,558
; PRIOR FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: US 60/160,587
; PRIOR FILING DATE: 1999-10-20
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 285
; LENGTH: 930
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-801-368-285

Alignment Scores:
Pred. No.: 3.2e-24 Length: 930
Score: 264.50 Matches: 71
Percent Similarity: 49.15% Conservative: 45
Best Local Similarity: 30.08% Mismatches: 71
Query Match: 18.24% Indels: 49
DB: 10 Gaps: 5

US-09-709-103-1F1 (1-282) x US-09-801-368-285 (1-930)
QY 25 TyrArgMetValIleLeuGlySerSerLysValGlyLysThrAlaIleValSerArgPhe 44
Db 31 TATAAGATAGTAGTTGTCGGTGGAGTGGCGTGGTAAATCTGCTTTAAACAATTCATTC 90
QY 45 LeuThrGlyArgPheGluAspAlaTyrThrProThrIleGluAspPheHisArgLysPhe 64
Db 91 ATTCATCATCTTGTGGAGCAATATGACCTACTATCGAAGATCTTACAGAAACAA 150
QY 65 TyrSerIleArgGlyGluValTyrGlnLeuAspIleLeuAspThrSerGlyAsnHisPro 84
Db 151 GTTGTCATCGATGACAAAGATATCATTTTGACATCTAGATATCTGTGGCAAGAGAG 210
QY 85 PheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIleLeuValPheSer 104
Db 211 TATTCTGGATGAGAGACAGTACATGAGGACTGGGGAAGTTTCTACTGCTATTTC 270
QY 105 LeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGlnGlnIleLeuAspThr 124

```

```

Db 271 GTCACCTCTAGAAATCTCTTGTATGATGAGTACTGCTTATTATCAGCAATTT- 321
QY 125 LysSerCysLeuLysAsnLysThrLysGluAsnValAspValProLeuValIleCysGly 144
Db 322 -----CAAAGAGATAAAGATTTCTGACTACATTTCTGTAGTCGTGGTAGGT 366
QY 145 AsnLysGlyAspArgAspPheTyrArgGluValAspGln-----ArgGluIle 160
Db 367 AACAAATTTGGACCTTTGAAATGAAGACAAGTCTCTTATGAAGACGGGTACGCTGGCC 426
QY 161 GluGlnLeuValGlyAspProGlnArgCysAlaTyrPheGluIleSerAlaLysLys 180
Db 427 AAGCAGTTG-----AATGCACCTTTCTAGAAACGCTCTGCGAACAACAA 468
QY 181 AsnSerSerLeuAspGlnMetPheArgAlaLeuPheAlaMet----- 194
Db 469 GCCATCAACGTAGACGAGGCCCTTTTATAGCCTTATTCTGTTGGTAAGGACGACGGTGGG 528
QY 195 -----AlaLysLeu 197
Db 529 AAATACAATAGCATGAATCGTCAACTGGATTAATAGCAATGAATAAGAGATTCGAGCTA 588
QY 198 ProSerGluMetSerProAspLeuHisArgLysValSerValGlnTyrCysAspValLeu 217
Db 589 ACCTCATCTGCACACGCGATAGAGAAAAAAGAACACACGGGCTCTAT----- 636
QY 218 HisLysLysAlaLeuArgAsnLysLysLeuLeuArgAlaGlySerGly 233
Db 637 -----GTACTCGATAATTTCTTGACCAATGCTGCGCACTGGC 672

RESULT 14
US-09-867-701-5566
; Sequence 5566, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
; APPLICANT: Aglate, Paul A.
; APPLICANT: Jones, Robert
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; CURRENT FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5566
; LENGTH: 536
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-867-701-5566

Alignment Scores:
Pred. No.: 5.93e-24 Length: 536
Score: 259.50 Matches: 55
Percent Similarity: 60.43% Conservative: 29
Best Local Similarity: 39.57% Mismatches: 46
Query Match: 17.90% Indels: 9
DB: 10 Gaps: 2

US-09-709-103-1F1 (1-282) x US-09-867-701-5566 (1-536)
QY 25 TyrArgMetValIleLeuGlySerSerLysValGlyLysThrAlaIleValSerArgPhe 44
Db 131 TATAAACTTGTGTAGTTGGAGCTGTCGCTAGGCAAGAGTGCCTTCAGCATACAGCTA 190
QY 45 LeuThrGlyArgPheGluAspAlaTyrThrProThrIleGluAspPheHisArgLysPhe 64
Db 191 ATTCAGATCATTTTGTGGACGAATATGATCAACAATAGAGGATTCCTACAGGAGCAAC 250
QY 65 TyrSerIleArgGlyGluValTyrGlnLeuAspIleLeuAspThrSerGlyAsnHisPro 84
Db 251 GTAGTAATTGATGGAGAAACCTGCTCTCTGGATATTCTCGACACAGCAGGTCAAGAGGAG 310

```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 30, 2002, 14:39:31 / Search time 77 Seconds

(without alignments)
1123.154 Million cell updates/sec

Title: US-09-709-103-1f1

Perfect score: 1450
Sequence: 1 MKLAMIKMCPDSELSIP.....REKASAGSQAKDKRCVTS, 282

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame.p2n.model -DBV=xlp
-Q=/cgn2_1/uspto/TRANSDUS09709103/runat_30122002_143926_20678/app_query.fasta_1.455
-DB=Issued Patents NA -QFMT=fastap -SUFIX=p2n.m1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=biosum62 -TRANS=human40.cdi
-LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTMT=plo -NORMT=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=TRANSDUS09709103@cgn2_1.1.31@runat_30122002_143926_20678 -NCPUS=6 -ICPU=3
-NO XLPXY -NO WMAP -LARGEOUTERY -NEG SCORES=0 -WAIT -LONGLOG -DBV TIMEOUT=120
-WARN TIMEOUT=30 -TREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database:

Issued Patents NA:
1: /cgn2_6/prodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/prodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/prodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/prodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/prodata/1/ina/PCUS_COMB.seq.*
6: /cgn2_6/prodata/1/ina/backfilseq1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1449	99.9	1841	4	US-09-053-374A-1
2	1399.5	96.5	1689	4	US-09-053-374A-4
3	1387	95.7	3986	4	US-09-053-374A-3
4	1343	92.6	3079	4	US-09-053-374A-6
5	300	20.7	615	1	US-08-247-946A-5
6	300	20.7	615	1	PCT-US95-06420-5
7	279	19.2	5775	5	US-08-306-691B-15
8	279	19.2	5775	5	PCT-US93-06251-29
9	277.5	19.1	570	4	US-08-884-866A-2
10	277.5	19.1	570	4	US-08-884-866A-11
11	276	19.0	607	4	US-08-429-964-85
12	275.5	19.0	480	4	US-08-884-866A-9

13	271.5	18.7	4480	4	US-09-167-322-12	Sequence 12, Appl
14	266.5	18.4	450	4	US-08-884-866A-10	Sequence 10, Appl
15	266.5	18.4	574	2	US-08-429-964-83	Sequence 83, Appl
16	265.5	18.3	2436	1	US-08-306-691B-16	Sequence 16, Appl
17	237	16.3	2309	3	US-09-078-317-3	Sequence 3, Appl
18	237	16.3	2309	4	US-09-454-818-3	Sequence 3, Appl
19	234.5	16.2	600	3	US-09-078-317-1	Sequence 1, Appl
20	234.5	16.2	600	4	US-09-454-818-1	Sequence 1, Appl
21	219.5	15.1	563	4	US-09-385-982-426	Sequence 426, App
22	204	14.1	3497	4	US-09-503-505A-2	Sequence 2, Appl
23	186	12.8	6453	1	US-08-306-691B-14	Sequence 10, Appl
24	186	12.8	6453	3	US-09-209-668-10	Sequence 10, Appl
25	186	12.8	6453	3	US-09-356-952-8	Sequence 8, Appl
26	185.5	12.8	2964	2	US-08-846-790A-2	Sequence 2, Appl
27	185.5	12.8	2964	3	US-08-935-333-2	Sequence 2, Appl
28	180	12.4	702	3	US-08-842-976-2	Sequence 2, Appl
29	180	12.4	702	3	US-09-213-397-2	Sequence 2, Appl
30	180	12.4	702	3	US-09-416-489-2	Sequence 2, Appl
31	177.5	12.2	1166	5	PCT-US96-12129B-1	Sequence 1, Appl
32	175.5	12.1	603	4	US-09-325-932A-29	Sequence 29, Appl
33	175.5	12.1	932	4	US-09-325-932A-28	Sequence 28, Appl
34	175	12.1	1175	4	US-09-387-341-215	Sequence 215, App
35	173	11.9	1175	2	US-08-773-423-6	Sequence 6, Appl
36	172	11.9	897	2	US-09-006-535-2	Sequence 2, Appl
37	172	11.9	1443	1	US-08-076-089-1	Sequence 1, Appl
38	172	11.9	1443	2	US-08-707-200-1	Sequence 1, Appl
39	172	11.9	1443	4	US-08-996-565-1	Sequence 1, Appl
40	172	11.9	1443	5	PCT-US93-05643-1	Sequence 1, Appl
41	172	11.9	1525	2	US-09-006-535-1	Sequence 1, Appl
42	167	11.5	897	2	US-09-006-535-7	Sequence 7, Appl
43	166.5	11.5	5197	4	US-09-293-170-6	Sequence 6, Appl
44	166	11.4	890	4	US-09-075-454-14	Sequence 14, Appl
45	165	11.4	240	4	US-08-884-866A-8	Sequence 8, Appl

ALIGNMENTS

RESULT 1
US-09-053-374A-1
; Sequence 1, Application US/09053374A
; Patent No. 6462177
; GENERAL INFORMATION:
; APPLICANT: YEN, KWANG-MU
; TITLE OF INVENTION: MAMMALIAN BLOOD LOSS-INDUCED GENE, KD312
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: AMGEN INC.
; STREET: ONE AMGEN CENTER DRIVE
; CITY: THOUSAND OAKS
; STATE: CA
; COUNTRY: US
; ZIP: 91320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/053,374A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: COOK, ROBERT R.
; REGISTRATION NUMBER: 31,602
; REFERENCE/DOCKET NUMBER: A-514
; INFORMATION FOR SEQ ID NO. 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1841 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:

NAME/KEY: CDS
LOCATION: 255...1097
US-09-053-374A-1

Alignment Scores:

Pred. No.: 5,33e-166 Length: 1841
Score: 1449.00 Matches: 281
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.93% Indels: 0
DB: 4 Gaps: 0

US-09-709-103-1f1 (1-282) x US-09-053-374A-1 (1-1841)

QY 1 MetLysLeuAlaMetIleLysLysMetCysProSerAspSerGluLeuSerIlePro 20
DB 255 ATGAACACTGGCGCGATGATCAAGAGATGTGCCGAGCGACTCGGAGTATCCCG 314
QY 21 AlaLysAsnCysTyrArgMetValIleLeuGlySerLysValGlyLysThrAlaIle 40
DB 315 GCCAAGAACTGCTATCGCATGTCATCTCGGCTGCTCCAGGTGGGCAAGCGCCATC 374
QY 41 ValSerArgPheLeuThrGlyArgPheGluAspAlaTyrThrProThrIleGluAspPhe 60
DB 375 GTGTCGGCTTCTCACCAGCGCTTCGAGGACGCTACAGCCTTACCATCGAGACTTC 434
QY 61 HisArgLysPheTyrSerIleArgGlyGluValTyrGlnLeuAspIleLeuAspThrSer 80
DB 435 CACCGCAAGTTCTCTCATCGCGCGAGGTCTACCGACTCGACATCTCGACACGTCC 494
QY 81 GlyAsnHisProPheProAlaMetArgArgLeuSerIleLeuThrGlyValPheIle 100
DB 495 GGCAACACCCGCTTCCCGCATCGCGCTCTCATCTCTCACAGGAGAGCTTTTCATC 554
QY 101 LeuValPheSerLeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGlnGln 120
DB 555 CTGGTGTTCAGTCTGACAAACCGGACCTCTTCGAGGAGGTGCGAGCGCTCAGGCGAG 614
QY 121 IleLeuAspThrLysSerCysLeuLysAsnLysThrLysGluAsnValAspValProLeu 140
DB 615 ATCTCTGACACCAAGTCTTGCTCAGAACAAACCAAGGAGAACGTGGAGCTGCCCGCTG 674
QY 141 ValIleCysGlyAsnLysGlyAspArgAspPheTyrArgGluValAspGlnArgGluIle 160
DB 675 GTCTATCTCGGCAACAAAGGGTGACCGGACTTCTACCGCGAGGTGGACCGCGAGATC 734
QY 161 GluGlnLeuValGlyAspAspProGlnArgCysAlaTyrPheGluIleSerAlaLysLys 180
DB 735 GAGCAGCTGGTGGGCGACGACCCCGCGCTGCGCTTCTCGAGATCTCGGCCAAGAG 794
QY 181 AsnSerSerLeuAspGlnMetPheArgAlaLeuPheAlaMetAlaLysLeuProSerGlu 200
DB 795 AACAGCAGCTTGGACCATGTTTCGCGCGCTCTTCGCCATGGCCAGAGTTCGCCGCGAG 854
QY 201 MetSerProAspLeuHisArgLysValSerValGlnTyrCysAspValLeuHisLysLys 220
DB 855 ATGAGCCAGACCTGACCCGCAAGTCTCGGTGCGTACTGCGACGCTGCTGCACAGAG 914
QY 221 AlaLeuArgAsnLysLysLeuLeuArgAlaGlySerGlyGlyGlyGlyAspProGly 240
DB 915 GCGTTCGCGAACAAGAGCTCTCGGCGCGGCGAGCGCGCGCGCGCGCGCGCGCGCG 974
QY 241 AspAlaPheGlyIleValAlaProPheAlaArgArgProSerValHisSerAspLeuMet 260
DB 975 GACGCTTTGGCATCGTGGCACTTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1034
QY 261 TyrIleArgGluLysAlaSerAlaGlySerGlnAlaLysAspLysGluArgCysValIle 280
DB 1035 TACATCGCGAGAGGCG 1094
QY 281 Ser 281
DB 1095 AGC 1097

RESULT 2

US-09-053-374A-4
; Sequence 4, Application US/09053374A
; Patent No. 6462177
; GENERAL INFORMATION:
; APPLICANT: YEN, KWANG-MU
; TITLE OF INVENTION: MAMMALIAN BLOOD LOSS-INDUCED GENE, KD312
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: AMGEN INC.
; STREET: ONE AMGEN CENTER DRIVE
; CITY: THOUSAND OAKS
; STATE: CA
; COUNTRY: US
; ZIP: 91320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/053,374A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: COOK, ROBERT R.
; REGISTRATION NUMBER: 31,602
; REFERENCE/DOCKET NUMBER: A-514
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1689 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 132..971
; US-09-053-374A-4

Alignment Scores:

Pred. No.: 4,74e-160 Length: 1689
Score: 1399.50 Matches: 273
Percent Similarity: 98.22% Conservative: 3
Best Local Similarity: 97.15% Mismatches: 4
Query Match: 96.52% Indels: 1
DB: 4 Gaps: 1

US-09-709-103-1f1 (1-282) x US-09-053-374A-4 (1-1689)

QY 1 MetLysLeuAlaAlaMetIleLysLysMetCysProSerAspSerGluLeuSerIlePro 20
DB 132 ATGAACACTGGCGCGATGATCAAGAGATGTGCCAAGCGACTCTGAACTGAGTATCCCG 191
QY 21 AlaLysAsnCysTyrArgMetValIleLeuGlySerLysValGlyLysThrAlaIle 40
DB 192 GCCAAGAACTGCTCAGGATGTCTCTCGGCTCATCCAAAGTGGGCAAGCGCCATC 251
QY 41 ValSerArgPheLeuThrGlyArgPheGluAspAlaTyrThrProThrIleGluAspPhe 60
DB 252 GTGTCGCGCTTCTCAGCGCGCTTCGAGGACGCTTACACCCCTACCATTTGAGACTTC 311
QY 61 HisArgLysPheTyrSerIleArgGlyGluValTyrGlnLeuAspIleLeuAspThrSer 80
DB 312 CACCGAAAAGTTTACTCGATCCGCGCGAGTCTACCAAGTGGACATCTGCACACATCT 371
QY 81 GlyAsnHisProPheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIle 100
DB 372 GGCATCATCTGCTTCCCGCGCATCGCGCGCTCTCTATCTCTCAGGAGAGCTTTTCAT 431
QY 101 LeuValPheSerLeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGlnGln 120

```

Db 432 CTGGTGTTCAGCTTAGACAAACCGCAGCTCTTCGAGAGGTGCAAGGCTTCAACAGCAG 491
Qy 121 TleleuAspThlySerCyseuLeuysAsnlyThrlsGluAsnValAspValProleu 140
Db 492 ATCTTAGACACCAAGCTCTGTCTCAAGAACCAACCAAGAAATGTGACGTCCGCTG 551
Qy 141 ValileCyseGlyAsnlyGlyAspArgAspPheTyArgGluValAspGlnArgGluile 160
Db 552 GTCATTTGGCGGTACCAAGGGGACCGGAGCTTCTACCGCGAAGTGGAGCAGCGGAGATT 611
Qy 161 GluGlnLeuValGlyAspAspProGlnArgCysAlaTyrrPheGluileSerAlaLysLys 180
Db 612 GAGCACTGTGTGGCATATACCTCAGCGTGTGCTTCTTGAGATTTCGCGCAAGAG 671
Qy 181 AsnSerSerLeuAspGlnMetPheArgAlaLeuPheAlaMetAlaLysLeuProSerGlu 200
Db 672 AATAGAGCTTGAGCAAGATGTTCCGTGCGCTCTTTGCATGCGCAAGCTGCTAGCAG 731
Qy 201 MetSerProAspLeuHisArgLysValSerValGlnTyrrCysAspValLeuHisLysLys 220
Db 732 ATGAGCGCTTACTTGACCGCAAGGTGTGTGCTGCTGAGTGTGAGCGTGTGCAAAAAG 791
Qy 221 AlaLeuArgAsnlyLysLysLeuLeuArgAlaGlySerGlyGlyGlyGlyValAspProGly 240
Db 792 GCTCTGAGAACAGAGAGCTTGTGCGGCGGAGC---GAGGTGGGGGCGACACGGA 848
Qy 241 AspAlaPheGlyileValAlaProPheAlaArgArgProSerValHisSerAspLeuMet 260
Db 849 GATGCTTGGCATCTTGCGGCTTGTGCTGCGAGACTAGCGTGTGATGACACCTCATG 908
Qy 261 TyrlleArgGluLysAlaSerAlaGlySerGlnAlaLysAspLysGluArgCysValile 280
Db 909 TACATTCGTGAGAAACCAAGTGTGACGACGACGAGCTTAAGAGAGCGCTGTGTATC 968
Qy 281 Ser 281
Db 969 AGT 971

RESULT 3
US-09-053-374A-3
Sequence 3, Application US/09053374A
Patent No. 6462177
GENERAL INFORMATION:
APPLICANT: YEN, KWANG-MU
TITLE OF INVENTION: MAMMALIAN BLOOD LOSS-INDUCED GENE, KD312
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: AMGEN INC.
STREET: ONE AMGEN CENTER DRIVE
CITY: THOUSAND OAKS
STATE: CA
COUNTRY: US
ZIP: 91320
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/053.374A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: COOK, ROBERT R.
REGISTRATION NUMBER: 31,602
REFERENCE/DOCKET NUMBER: A-514
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3986 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)

```

```

US-09-053-374A-3
Alignment Scores:
Pred. No.: 5.53e-158 Length: 3986
Score: 1387.00 Matches: 281
Percent Similarity: 79.83% Conservative: 0
Best Local Similarity: 79.83% Mismatches: 0
Query Match: 95.66% Indels: 71
DB: Gaps: 1

US-09-709-103-1f1 (1-282) x US-09-053-374A-3 (1-3986)
Qy 1 MetLysLeuAlaAlaMetLileLysLysMetCysProSerAspSerGluLeuSerLlePro 20
Db 774 ATGAAACTGGCCGCGCATGATCAAGAAAGATGTGCCGAGCATCTGGAGCTGAGATCCCG 833
Qy 21 AlaLysAsnCyseTyrrArgMetValileLengLysSerLysValGlyLysThrAlaile 40
Db 834 GCCAAGAACCTGCTATCGCATGCTCATCTCTGCTGCTCAAGTGGGCAAGCCGCAATC 893
Qy 41 ValSerArgPheLeuThrGlyArgPheGluAspAlaTyrrThrProThrileGluAspPhe 60
Db 894 GTGTGCGCTTCTCTCACCGCGCGCTTCGAGAGCGCTTACAGCCTTACATGAGACTTC 953
Qy 61 HisArgLysPheTyserileArgGlyGluValTyrrGlnLeuAspLileLeuAspThrSer 80
Db 954 CACCGAAGTTCTATCCATCCGCGCGGAGGTCTACAGCTCGACATCTTGACAGCTTC 1013
Qy 81 GlyAsnHisProPheProAlaMetArgArgLeuSerileLeuThr----- 95
Db 1014 GGCACACACCGGTTCCCGCAGGCGGCGCTTCTCATCTCAGAGTGAAGCGGGGCGC 1073
Qy 95 ----- 95
Db 1074 GGGAGGTGGGGAGGGAAGGCGGGGAACTTCGCGCAGGCGCGCGCGCGCGT 1133
Qy 95 ----- 95
Db 1134 CCGGCTGCTGCGCGCGAGTAGTGCGCTTCGCGCTTAGAGAGCTAGCGCGCGCGCG 1193
Qy 95 ----- 95
Db 1194 GCTTCAAAATGACCGCGCATTTGCTCCCTGGCGCGCACCTCACCTTCTCTTTCTGCT 1253
Qy 96 -----GlyAspValPheLileLeuValPheSerLeuAspAsnArgAspSe 110
Db 1254 CTCTGTGCCCCCTTAGAGAGAGCTTTTCATCTCGGTGTTCAGTCTGAGACACGCGACTC 1313
Qy 110 rPheGluGluValGlnArgLeuArgGlnLileLeuAspThrLysSerCysLeuLysAs 130
Db 1314 CTTGAGAGAGGTGACGCGCTCGAGGCGAGATCTCGACACCAAGCTTTGCTCAAGAA 1373
Qy 130 nlySerThrLysGluAsnValAspValProLeuValileCyseGlyAsnlyGlyAspArgAs 150
Db 1374 CAAGAACCAAGAGAAAGTGGAGCGTCCCTGCTCATCTGCGGCAACAGGGTGAACGCGCA 1433
Qy 150 rPheTyrrArgGluValAspGlnArgGluileGluGlnLeuValGlyAspAspProGlnAr 170
Db 1434 CTTTACCGCGAGGTGAGCGCAGCGCGAGATCGACACACTGTGGGCGACGACCGAGCG 1493
Qy 170 gCyseAlaTyrrPheGluileSerAlaLysLysAsnSerSerLeuAspGlnMetPheArgAl 190
Db 1494 CTGGCGCTTACTTGAAGTCTGCGCCAGAGAAACAGCAGCTTGAACCAAGTGTCCGCGC 1553
Qy 190 AleuPheAlaMetAlaLysLeuProSerGlnMetSerProAspLeuHisArgLysValSe 210
Db 1554 GCTCTTGGCATGCGCAAGCTTCCAGCGAGATGAGCGCAACCTTGAACCGCAAGGCTTC 1613
Qy 210 rValGlnTyrrCysAspValLeuHisLysLysLysLysLysLysLysLysLysLysLys 230
Db 1614 GGTGAGTACTCGACGTCGTGCAAGAGGCGCTGGGAAACAAGAGCTGCTGGGCGG 1673
Qy 230 agLysSerGlyGlyGlyGlyGlyAspProGlyAspAlaPheGlyileValAlaProPheAl 250

```



```

COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/247,946A
FILING DATE: 24-MAY-1994
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: DOROTHY R. AUTH
REGISTRATION NUMBER: 36,434
REFERENCE/DOCKET NUMBER: 2026-4150
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 615
TYPE: Nucleic acid
STRANDEDNESS: Double
TOPOLOGY: Unknown
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Human
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPOTYPE:
TISSUE TYPE:
CELL TYPE:
CELL LINE:
ORGANELLE:
FEATURE:
NAME/KEY: TC21 gene
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION:
US-08-247-946A-5

Alignment Scores:
Pred. No.: 2,45e-27 Length: 615
Score: 300.00 Matches: 71
Percent Similarity: 55.32% Conservatve: 33
Best Local Similarity: 37.77% Mismatches: 70
Query Match: 20.69% Indels: 14
DB: 1 Gaps: 3

US-09-709-103-1f1 (1-282) x US-08-247-946A-5 (1-615)
QY 25 TyArgMetValIleLeuGlySerSerIleValGlyIleThrAlaIleValSerArgPhe 44
Db 43 TACCGGCTCGGTGGTGGCGGGGGCGGCGTGGCAAGTCGGCGTCCACATCCAGTTC 102
QY 45 LeuThrGlyArgPheGluAspAlaTyrThrProThrIleGluAspPheHisArgIysPhe 64
Db 103 ATCCAGTCCATTGTTGAACGATTAATGATCCACATTAAGTTTTCACAAACAG 162
QY 65 TySerIleArgGlyGluValTyrGlnLeuAspIleLeuAspThrSerGlyAsnHisPro 84
Db 163 TGTGTATGATAGTACAGACAGCCCGGTAGATATTTGGATACAGACAGANNNGAAG 222
QY 85 PheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIleLeuValPheSer 104
Db 223 TTTGGAGCCATGAGAAACAGTATATGAGCTGCGAAGGCTTCGTGTGGTTTTC 282
QY 105 LeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGlnGlnIleLeuAspThr 124
Db 283 GTCAAGATAGAGGAGAGTTTGAAGAAATCTATAAGTTCAAGACAGATTCTC----- 336
QY 125 LysSerCysLeuLysAsnLysThrIleGluAsnValAspValProLeuValIleCysGly 144
Db 337 -----AGAGTAAAGATCGTGTATGAGTTCCCAATGATTATTAATGGT 378

```

```

QY 145 AsnLysGlyAspArgAspPheTyrArgGluValAspGlnArgGluIleGluGlnLeuVal 164
Db 379 AATTAAGCACATCTGATCATCAAAAGACAGGTAAACAGGAAGAAAGCAACAGTTAGCA 438
QY 165 GlyAspAspProGlnArgCysAlaTyrPheGluIleSerAlaLysLysAsnSerIleu 184
Db 439 CGGACG-----CTTAAGCTTAACATCATGAGCAGTCAGCAAAAGATTGATGAATGTA 492
QY 185 AspGlnMetPheArgAlaLeuPheAlaMetAlaLysLeuProSerGluMet----- 201
Db 493 GATCAAGCTTTCATGCACTGTCGGGTTATCAGGAATTTCAAGACAGCAATGCTCT 552
QY 202 ---SerProAspLeuHisArgIys 208
Db 553 CTTTACACAGAACCAACGCGAAA 576

RESULT 6
PCT-US95-06420-5
Sequence 5, Application PC/TUS9506420
GENERAL INFORMATION:
APPLICANT: AARONSON, S.A.; CHAN, A.;
APPLICANT: MIKI, T. NOVEL HUMAN P4S-RELATED
TITLE OF INVENTION: ONCOGENES UNMASKED BY EXPRESSION OF CDNA
TITLE OF INVENTION: CLONING
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154

COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/06420
FILING DATE: 24-MAY-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/247,946
FILING DATE: 24-MAY-1994
ATTORNEY/AGENT INFORMATION:
NAME: DOROTHY R. AUTH
REGISTRATION NUMBER: 36,434
REFERENCE/DOCKET NUMBER: 2026-4150PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 615
TYPE: Nucleic acid
STRANDEDNESS: Double
TOPOLOGY: Unknown
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Human
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPOTYPE:
TISSUE TYPE:
CELL TYPE:
CELL LINE:
ORGANELLE:
FEATURE:

```


NAME/KEY: TC21 gene
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION:
PCT-US95-06420-5

Alignment Scores:

Pred. No.: 2,456-27 Length: 615
Score: 300.00 Matches: 71
Percent Similarity: 55.32% Conservative: 33
Best Local Similarity: 37.77% Mismatches: 70
Query Match: 20.69% Indels: 14
DB: 5 Gaps: 3

US-09-709-103-1f1 (1-282) x PCT-US95-06420-5 (1-615)

QY 25 TyrArgMetValIleLeuGlySerLysValGlyThrAlaIleValSerArgPhe 44
Db 43 TACCGGCTCGTGTGTCGGCGGGCGCGTGGCAAGTCGCGCTCACCATCCAGTTC 102
QY 45 LeuThrGlyArgPheGluAspAlaTyrThrProThrIleGluAspPheHisArgLysPhe 64
Db 103 ATCCAGTCCTATTGTTGAACGGATTATGATCCAAACATTGAAGATTCCTACACAAAGCAG 162
QY 65 TyrSerIleArgGlyGluValTyrGlnLeuAspIleLeuAspThrSerGlyAsnHisPro 84
Db 163 TGTGTGATAGATCACAGACAGCCCGCTAGATATTTGGATACAGCAGGANNNGAAGAG 222
QY 85 PheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIleLeuValPheSer 104
Db 223 TTTGGAGCCATGAGAAACAGTATATAGGACTGCGGAAGGCTTCTGTGTGCTTTTCA 282
QY 105 LeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGlnGlnIleLeuAspThr 124
Db 283 GTACAGATAGAGGAGGATTTTGAAGAAATCTATAAGTTTCAAGACAGATTCCTC 336
QY 125 LysSerCysLeuLysAsnLysThrLysGluAsnValAspValProLeuValIleCysGly 144
Db 337 -----AGCTAAGGATCGTGATGAGTCCCAATGATTTAATGGT 378
QY 145 AsnLysGlyAspArgAspPheTyrArgGluValAspGlnArgGluIleGluLeuVal 164
Db 379 AATAAAGCAGATCTGGATCATCAAGACAGAGTAAACACAGGAAGAAGACACAGTTAGCA 438
QY 165 GlyAspAspProGlnArgCysAlaTyrPheGluIleSerAlaLysLysAsnSerSerLeu 184
Db 439 CGCGAG-----CTTAAGTAACATACATGAGGATCAGCAAGATTAGGATGAATGTA 492
QY 185 AspGlnMetPheArgAlaLeuPheAlaMetAlaLysLeuProSerGluMet----- 201
Db 493 GATCAAGCTTTCCATGAATTTGTCGGGTATATCAGGAATTTCAAGACAGCAAGATGTCCT 552
QY 202 ---SerProAspLeuHisArgLys 208
Db 553 COTTCACAGAACCAACACGGAAA 576

RESULT 7

US-08-306-691B-15

Sequence 15, Application US/08306691B

Patent No. 5734039

GENERAL INFORMATION:

APPLICANT: Calabretta, Bruno

APPLICANT: Skorski, Tomasz

TITLE OF INVENTION: ANTIGENSE

TITLE OF INVENTION: OLIGONUCLEOTIDES TARGETING COOPERATING ONCOGENES

NUMBER OF SEQUENCES: 55

CORRESPONDENCE ADDRESS:

ADDRESSEE: Seidel, Gonda, Lavorigna & Monaco, P.C.

STREET: Two Penn Center, Suite 1800

CITY: Philadelphia

STATE: Pennsylvania

COUNTRY: U.S.A.

ZIP: 19102

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WordPerfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/306,691B
FILING DATE: September 15, 1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:

ATTORNEY/AGENT INFORMATION:
NAME: Monaco, Daniel A.
REGISTRATION NUMBER: 30,480
REFERENCE/DOCKET NUMBER: 8321-8
TELEPHONE: (215) 568-8383
TELEFAX: (215) 568-5549
TELEX: No. 5734039e

INFORMATION FOR SEQ ID NO: 15:

SEQUENCE CHARACTERISTICS:

LENGTH: 5775 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

US-08-306-691B-15

Alignment Scores:

Pred. No.: 2,376-23 Length: 5775

Score: 279.00 Matches: 67

Percent Similarity: 54.59% Conservative: 34

Best Local Similarity: 36.22% Mismatches: 72

Query Match: 19.24% Indels: 12

DB: 1 Gaps: 4

US-09-709-103-1f1 (1-282) x US-08-306-691B-15 (1-5775)

QY 25 TyrArgMetValIleLeuGlySerLysValGlyThrAlaIleValSerArgPhe 44

Db 202 TATAAAGCTTGTGTGTCGGCTGCGTAGGACAGAGTTCCTGACGATACAGCTA 261

QY 45 LeuThrGlyArgPheGluAspAlaTyrThrProThrIleGluAspPheHisArgLysPhe 64

Db 262 ATTCAGAAATCATTTGTGTGACGAATATGATCAACAATAGAGGATTCCTACAGGAAGCAA 321

QY 65 TyrSerIleArgGlyGluValTyrGlnLeuAspIleLeuAspThrSerGlyAsnHisPro 84

Db 322 GTAGTAATTCGATGGAGAACCTGCTCTTTGGATATTCGACACAGCAGGTCAGAGGAG 381

QY 85 PheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIleLeuValPheSer 104

Db 382 TACAGTGCAATGAGGACCCAGTACATGAGGACTGGGAGGCTTCTTTGTGATTTGCC 441

QY 105 LeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGlnGlnIleLeuAspThr 124

Db 442 ATAAATAATACCTAAATCATTTGAAGATATTCACCATTTATAGAGAACAAAT 492

QY 125 LysSerCysLeuLysAsnLysThrLysGluAsnValAspValProLeuValIleCysGly 144

Db 493 -----AAAAGAGTTAAGGACTCTGAAGATGTACCTATGCTCTTAGTAGA 537

QY 145 AsnLysGlyAspArgAspPheTyrArgGluValAspGlnArgGluIleGluLeuVal 164

Db 538 AATAAATGTGATTTGCCT---TCTAGAACAGTAGACACAAACAGGCTCAGGACTTAGCA 594

QY 165 GlyAspAspProGlnArgCysAlaTyrPheGluIleSerAlaLysLysAsnSerSerLeu 184

Db 595 AGAAGT-----TATGGAATTCCTTTTATGAAACATCAGCAAAAGACAGGCTGTT 648

QY 185 AspGlnMetPheArgAlaLeuPhe---AlaMetAlaLysLeuProSerGluMetSerPro 203

Db 649 GATGATGCTTCTATACATTACTTGGAGAAATTCGAAAAACATAAGAAAGATGACAAA 708
 Qy 204 AspleuHieArglys 208
 Db 709 GATGGTAAAAAGAG 723

RESULT 8

PCT-US93-06251-29
 ; Sequence 29, Application PC/TUS9306251
 ; GENERAL INFORMATION:
 ; APPLICANT: Wickstrom, Eric and Rife, Jason P.
 ; TITLE OF INVENTION: Trivalent Synthesis of Oligonucleotides Containing
 ; TITLE OF INVENTION: Stereospecific Alkylphosphonates and Arylphosphonates
 ; NUMBER OF SEQUENCES: 93
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESS: SCULLY, SCOTT, MURPHY & PRESSER
 ; STREET: 400 Garden City Plaza
 ; CITY: Garden City
 ; STATE: NY
 ; COUNTRY: USA
 ; ZIP: 11530
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US93/06251
 ; FILING DATE: 19930630
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Digilio, Frank S.
 ; REGISTRATION NUMBER: 31,346
 ; REFERENCE/DOCKET NUMBER: 8586
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 516-742-4343
 ; TELEFAX: 516-742-4366
 ; TELEX: 230 901 SANS UR
 ; INFORMATION FOR SEQ ID NO: 29:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 5775 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; PCT-US93-06251-29

Alignment Scores:

Pred. No.: 2.37e-23 Length: 5775
 Score: 279.00 Matches: 67
 Percent Similarity: 54.59% Conservative: 34
 Best Local Similarity: 36.22% Mismatches: 72
 Query Match: 19.24% Indels: 12
 DB: 5 Gaps: 4

US-09-709-103-1f1 (1-282) x PCT-US93-06251-29 (1-5775)

Qy 25 TyrArgMetValIleLeuGlySerSerIysValGlyLysThrAlaIleValSerArgPhe 44
 Db 202 TATAAAGCTTGGTGGTGGAGCTTGGCGTAGGCAAGAGTCCCTTACCATTCAGCTA 261
 Qy 45 LeuThrGlyArgPheGluAspAlaTyrThrProThrIleGluAspPheHisArgLysPhe 64
 Db 262 ATTCGAATCATTTTGTGGAGCAATATGATCCACAAATAGAGGATTCCTACAGAGAGCA 321
 Qy 65 TyrSerIleArgGlyGluValTyrGlnLeuAspIleLeuAspThrSerGlyAsnHisPro 84
 Db 322 GTAGTAATTTGATGGAGAAACCTGCTCTTGGATATTCGACACAGCGGTCAAGAGAG 381
 Qy 85 PheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIleLeuValPheSer 104
 Db 382 TACACTGCATGAGGACCAAGTACATGAGACTGGGGAGGGCTTCTTGTGTATTGGCC 441

Qy 105 LeuAspAsnHisArgSerPheGluGluValGlnArgLeuArgGlnIleLeuAspThr 124
 Db 442 ATAAATATCTAATATCATTTGAGATATTCACCATTAATAGAGAACTAAT 492
 Qy 125 LysSerCysLeuIysAsnLysThrIysGluAsnValAspValProLeuValIleCysGly 144
 Db 493 -----AAAAGAGTTAAGACTGGAAGATGCTTATGCTTCTTCTAGTAGCA 537
 Qy 145 AsnLysGlyAspArgAspPheTyrArgGluValAspGlnArgGluIleGluIleuVal 164
 Db 538 AATAATATGATTTGGCT---TCTNACAAGTAGACACAAAACAGGCTCAGACCTTAGCA 594
 Qy 165 GlyAspAspProGlnArgCysAlaTyrPheGluIleSerAlaLysLysAsnSerIeu 184
 Db 595 AGAAGT-----TATGAAATTCCTTTATTGAAACATCAGCAAGACAGAGGTGT 648
 Qy 185 AspGlnMetPheArgAlaLeuPhe---AlaMetAlaLysLeuProSerGlnMetSerPro 203
 Db 649 GATGATGCTTCTATACATTAGTTGCGAATAATTCGAAACATTAAGAAAGATGAGCAAA 708
 Qy 204 AspleuHieArglys 208
 Db 709 GATGGTAAAAAGAG 723

RESULT 9

US-08-884-866A-2
 ; Sequence 2, Application US/0884866A
 ; GENERAL INFORMATION:
 ; APPLICANT: Chien, Shu
 ; APPLICANT: Shyy, John Y-J
 ; TITLE OF INVENTION: GENE THERAPY IN CORONARY ANGIOPLASTY AND
 ; FILE REFERENCE: BYPASS
 ; CURRENT APPLICATION NUMBER: US/08/884,866A
 ; CURRENT FILING DATE: 1997-06-30
 ; PRIOR APPLICATION NUMBER: 60/030,358
 ; PRIOR FILING DATE: 1996-11-08
 ; NUMBER OF SEQ ID NOS: 25
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 570
 ; TYPE: DNA
 ; ORGANISM: Rattus norvegicus
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)...(570)
 ; US-08-884-866A-2

Alignment Scores:

Pred. No.: 1.18e-24 Length: 570
 Score: 277.50 Matches: 73
 Percent Similarity: 53.23% Conservative: 34
 Best Local Similarity: 36.32% Mismatches: 72
 Query Match: 19.14% Indels: 23
 DB: 4 Gaps: 6

US-09-709-103-1f1 (1-282) x US-08-884-866A-2 (1-570)

Qy 25 TyrArgMetValIleLeuGlySerSerIysValGlyLysThrAlaIleValSerArgPhe 44
 Db 10 TATAAAGCTTGGTGGTGGAGCTTGGCGTAGGCAAGAGTCCCTTACCATTCAGCTA 69
 Qy 45 LeuThrGlyArgPheGluAspAlaTyrThrProThrIleGluAspPheHisArgLysPhe 64
 Db 70 ATCCAGAACCATTTTGTGACCAATACGACCCCACTATAGAGATTCCTACCGGAAGCAG 129
 Qy 65 TyrSerIleArgGlyGluValTyrGlnLeuAspIleLeuAspThrSerGlyAsnHisPro 84
 Db 130 GTGGTATTGATGGGAGACGTGCTGTGGACATCTGTAACCGCGGCTGGAGAG 189
 Qy 85 PheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIleLeuValPheSer 104
 Db 190 TACAGCGCATGCGGAGCAAGTCAATGCGACCGGGAGGGCTTCTGTGTATTGGCC 249

```

Qy 105 LeuAspAsnArgAspSerPheGluValGlnArgLeuArgGlnGlnLeuLeuAspThr 124
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 250 ATCAACAACCAAGTCTTTTGGAGACATCCACCAGTACAGGAGCAGATC----- 300

Qy 125 LysSerCysLeuLysAsnLysThrLysGluAsnValAspValProLeuValIleCysGly 144
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 301 -----AAACGGGTGAAGACTCGGATGACGTGCCATCGCTGGTGGTGGGG 345

Qy 145 AsnLysGlyAspArgAspPheTyArgGluValAspGlnArgGluIleGluGlnLeuVal 164
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 346 AACAAAGTGTGACCTG---GCTGCACGCACCTGTGGAATCTCGGCAGGCTCAGACCTCGCC 402

Qy 165 GlyAspAspProGlnArgCysAlaTyPheGluIleSerAlaLysLysAsnSerSerLeu 184
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 403 CGAAGC-----TACGGCATCCCTACATCGAGACCTCGGCCAAGACCCGGCAGGAGTG 456

Qy 185 AspGlnMetPheArgAlaLeu-----PheAlaMetAlaLysLeu--- 197
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 457 GAGGATGCTTCTACAGTGTGGTGGTGAGATCCGGCAGCACAAGCTGCGAAGCTGAAC 516

Qy 198 ---ProSerGluMetSerProAspLeuHisArgLysValSerValGlnTyrCysAspVal 216
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 517 CCTCTGATGAGAGTGGCCCGG-CTGCAT-----GAGCTGCAAGTGTGTCTCTC 566

Qy 217 Leu 217
   : : : : :
Db 567 CTG 569

RESULT 10
US-08-884-866A-11
; Sequence 11, Application US/08884866A
; GENERAL INFORMATION:
; APPLICANT: Chien, Shu
; APPLICANT: Shyy, John Y-J
; TITLE OF INVENTION: GENE THERAPY IN CORONARY ANGIOPLASTY AND
; FILE OF INVENTION: BYPASS
; FILE REFERENCE: UCSD1100-1
; CURRENT APPLICATION NUMBER: US/08/884,866A
; CURRENT FILING DATE: 1997-06-30
; PRIOR APPLICATION NUMBER: 60/030,358
; PRIOR FILING DATE: 1996-11-08
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 570
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (49)...(51)
; OTHER INFORMATION: nnn = Any nucleic acid triplet, except for UCA,
; OTHER INFORMATION: UCC, UCG, UCU, AGC, or AGU
; OTHER INFORMATION: Variation of SEQ ID NO.:2
US-08-884-866A-11

Alignment Scores:
Pred. No.: 118e-24 Length: 570
Score: 277.50 Matches: 73
Percent Similarity: 53.23% Conservative: 34
Best Local Similarity: 36.32% Mismatches: 72
Query Match: 19.14% Indels: 23
DB: 4 Gaps: 6

US-09-709-103-1f1 (1-282) x US-08-884-866A-11 (1-570)

Qy 25 TyrArgMetValIleLeuGlySerSerLysValGlyLysThrAlaIleValSerArgPhe 44
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 10 TATAAGCTGGTGGTGGCGCGCGGTGTGGCAANNNGCGCTGACCATCCAGCTG 69

Qy 45 LeuThrGlyArgPheGluAspAlaTyThrProThrIleGluAspPheHisArgLysPhe 64
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 70 ATCCAGAACCACTTTGTGAGCAATACGACCCCATATAGAGGATTCCTACCGAAGCAG 129

```

```

Qy 65 TyrSerIleArgGlyGluValTyrGlnLeuAspIleLeuAspThrSerGlyAsnHisPro 84
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 130 GTGGTCAATGTATGGGAGACGCTGCTGTGGACATCTCTGGATACCGCGGCTCGAGAG 189

Qy 85 PheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIleLeuValPheSer 104
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 190 TACACGGCATGCGGACCATGATGCGCACCGGGAGGGCTTCTCTGTGTGTGTTGCC 249

Qy 105 LeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGlnGlnIleLeuAspThr 124
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 250 ATCAACAACCAAGTCTTTTGGAGACATCCACCAGTACAGGAGCAGATC----- 300

Qy 125 LysSerCysLeuLysAsnLysThrLysGluAsnValAspValProLeuValIleCysGly 144
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 301 -----AAACGGGTGAAGACTCGGATGACGTGCCATCGCTGGTGGTGGGG 345

Qy 145 AsnLysGlyAspArgAspPheTyArgGluValAspGlnArgGluIleGluGlnLeuVal 164
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 346 AACAAAGTGTGACCTG---GCTGCACGCACCTGTGGAATCTCGGCAGGCTCAGACCTCGCC 402

Qy 165 GlyAspAspProGlnArgCysAlaTyPheGluIleSerAlaLysLysAsnSerSerLeu 184
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 403 CGAAGC-----TACGGCATCCCTACATCGAGACCTCGGCCAAGACCCGGCAGGAGTG 456

Qy 185 AspGlnMetPheArgAlaLeu-----PheAlaMetAlaLysLeu--- 197
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 457 GAGGATGCTTCTACAGTGTGGTGGTGAGATCCGGCAGCACAAGCTGCGAAGCTGAAC 516

Qy 198 ---ProSerGluMetSerProAspLeuHisArgLysValSerValGlnTyrCysAspVal 216
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 517 CCTCTGATGAGAGTGGCCCGG-CTGCAT-----GAGCTGCAAGTGTGTCTCTC 566

Qy 217 Leu 217
   : : : : :
Db 567 CTG 569

RESULT 11
US-08-429-964-85
; Sequence 85, Application US/08429964
; Patent No. 5962243
; GENERAL INFORMATION:
; APPLICANT: BROWN, MICHAEL S.
; APPLICANT: GOLDSTEIN, JOSEPH L.
; APPLICANT: REISS, YUVAL
; APPLICANT: JAMES, GUY L.
; TITLE OF INVENTION: METHODS FOR THE IDENTIFICATION OF FARNESYL
; TITLE OF INVENTION: TRANSFERASE INHIBITORS
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ARNOLD, WHITE & DURKEE
; STREET: P.O. BOX 4433
; CITY: HOUSTON
; STATE: TEXAS
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/08/429,964
; FILING DATE: 27-APR-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/021,625
; FILING DATE: 16-FEB-1993
; CLASSIFICATION: 435
; APPLICATION NUMBER: US 07/822,011
; FILING DATE: ABANDONED
; CLASSIFICATION: 435
; APPLICATION NUMBER: PCT/US/91/02650

```

FILING DATE: 18-APR-1991
 CLASSIFICATION: 435
 APPLICATION NUMBER: US 07/615,715
 FILING DATE: 20-NOV-1990
 CLASSIFICATION: 435
 APPLICATION NUMBER: 435
 FILING DATE: 18-APR-1990 (ABANDONED)
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: PARKER, DAVID L.
 REGISTRATION NUMBER: 32,165
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (512) 418-3000
 TELEFAX: (713) 789-2679
 TELETYPE: 79-0924
 INFORMATION FOR SEQ ID NO: 85:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 607 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-08-429-964-85

Alignment Scores:
 Pred. No.: 1.96e-24 Length: 607
 Score: 276.00 Matches: 66
 Percent Similarity: 54.59% Conservative: 35
 Best Local Similarity: 35.68% Mismatches: 72
 Query Match: 19.03% Indels: 12
 Gaps: 4

US-09-709-103-1f1 (1-282) x US-08-429-964-85 (1-607)

QY 25 TYRAGMETHALLEUENGLYSESERLYSVALGLYVSTHRALEVALSERARGPHE 44
 DB 10 TATTAAGCTGTGTGGTGGAGCTGTGGCGTAGCGAAGAGCTTACGATACGATCA 69
 QY 45 LEUTHRGVARGPHEGLUASPALATYRTHRPROTHRILEGLUASPPHEHISARGLYS 64
 DB 70 ATTGACATCATTTTGTGGACGATATGATCCACATATAGAGATTCCTACGAGAGCA 129
 QY 65 TYRSEIRLEARGGLYGLUVALTYRGLINLEUASPILEUASPTHSERGLYASNHS 84
 DB 130 GTAGTAATTTGAGGAGAACTGTCTGTGATATTCGACACAGCGTCAAGAGAG 189
 QY 85 PHEPROLAMETARGARGLEUSERILEUTHRGVARGPVALPHEILEUVALPHESE 104
 DB 190 TACAGTGCATAGGAGACCACTGATGAGAGCTGGGAGGCTTCTGTGATTTGCC 249
 QY 105 LEUASPASNARGASPSERPHEGLUVALGINARGLEUARGGLINGLILEUASPT 124
 DB 250 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 300
 QY 125 LYSSERCYSEULYSASNULYSTRYLSEGLUASNVALASPVALPROLEUVALILECYSG 144
 DB 301 -----AAAAGAGTTAGAGACTCTGAAAGTATGATGCTCTGATAGGA 345
 QY 145 ASNLYSGLYASPARGASPPHETRYARGGLUVALASPGLINARGGLULEGLUINLEUVAL 164
 DB 346 AATTAATGTGATTTGCCCTTCTTAAGAACAGTACCAAAACAGGCTCAAGACTTACA 402
 QY 165 GLYASPARPROGLINARGCYSAATYRPHGGLULESERALALYLYSASNSE 184
 DB 403 AGAAGT-----TATGGAATTCCTTTATTCACACATAGCAAGAGAGAGGCTT 456
 QY 185 ASPLMETPHEARGALALEUPHE---ALAMETALYLYSLEUPROSEGLIMETSE 203
 DB 457 GATGATGCCCTTCTATACATTAAGTTCGAGAAATTCGAAAGATTAAGAAAGATGACGAA 516
 QY 204 ASPLMETPHEARGALALEU 208
 DB 517 GATGATGCTTCTTACACGTTG 531

RESULT 12
 US-08-884-866A-9
 Sequence 9, Application US/08884866A
 GENERAL INFORMATION:
 APPLICANT: Chien, Shu
 TITLE OF INVENTION: GENE THERAPY IN CORONARY ANGIOPLASTY AND
 TITLE OF INVENTION: BYPASS
 FILE REFERENCE: US08100-1
 CURRENT APPLICATION NUMBER: US/08/884,866A
 CURRENT FILING DATE: 1997-06-30
 PRIOR APPLICATION NUMBER: 60/030,358
 NUMBER OF SEQ ID NOS: 25
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 9
 LENGTH: 480
 TYPE: DNA
 ORGANISM: Rattus norvegicus
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (49)...(51)
 OTHER INFORMATION: mm = Any nucleic acid triplet, except for UCA,
 OTHER INFORMATION: UCC, UCG, UCU, AGC, or AGU
 US-08-884-866A-9

Alignment Scores:
 Pred. No.: 1.59e-24 Length: 480
 Score: 275.50 Matches: 63
 Percent Similarity: 56.29% Conservative: 31
 Best Local Similarity: 37.72% Mismatches: 62
 Query Match: 19.00% Indels: 11
 Gaps: 3

US-09-709-103-1f1 (1-282) x US-08-884-866A-9 (1-480)

QY 25 TYRAGMETHALLEUENGLYSESERLYSVALGLYVSTHRALEVALSERARGPHE 44
 DB 10 TATTAAGCTGTGTGGTGGAGCTGTGGCGTAGCGAAGAGCTTACGATACGATCA 69
 QY 45 LEUTHRGVARGPHEGLUASPALATYRTHRPROTHRILEGLUASPPHEHISARGLYS 64
 DB 70 ATTGACATCATTTTGTGGACGATATGATCCACATATAGAGATTCCTACGAGAGCA 129
 QY 65 TYRSEIRLEARGGLYGLUVALTYRGLINLEUASPILEUASPTHSERGLYASNHS 84
 DB 130 GTAGTAATTTGAGGAGAACTGTCTGTGATATTCGACACAGCGTCAAGAGAG 189
 QY 85 PHEPROLAMETARGARGLEUSERILEUTHRGVARGPVALPHEILEUVALPHESE 104
 DB 190 TACAGTGCATAGGAGACCACTGATGAGAGCTGGGAGGCTTCTGTGATTTGCC 249
 QY 105 LEUASPASNARGASPSERPHEGLUVALGINARGLEUARGGLINGLILEUASPT 124
 DB 250 ATCAACACACCAAGCTTTTGAAGACATCCACAGTACAGGAGAGAGATTC----- 300
 QY 125 LYSSERCYSEULYSASNULYSTRYLSEGLUASNVALASPVALPROLEUVALILECYSG 144
 DB 301 -----AAACGGGTGAAGAGCTCGATGAGTCCCATGTGCTGTGGGG 345
 QY 145 ASNLYSGLYASPARGASPPHETRYARGGLUVALASPGLINARGGLULEGLUINLEUVAL 164
 DB 346 AACAAAGTGTGACTG---GCTGACGACACTGTGGAATTCGCGAGGCTTACGAGACTCGCC 402
 QY 165 GLYASPARPROGLINARGCYSAATYRPHGGLULESERALALYLYSASNSE 184
 DB 403 CGAAGC-----TACGCAATCCCTTACATCAGAGACCTCGGCAAGACCGGAGAGAGTG 456
 QY 185 ASPLMETPHEARGALALEU 191
 DB 457 GAGGATGCCCTTCTTACACGTTG 477

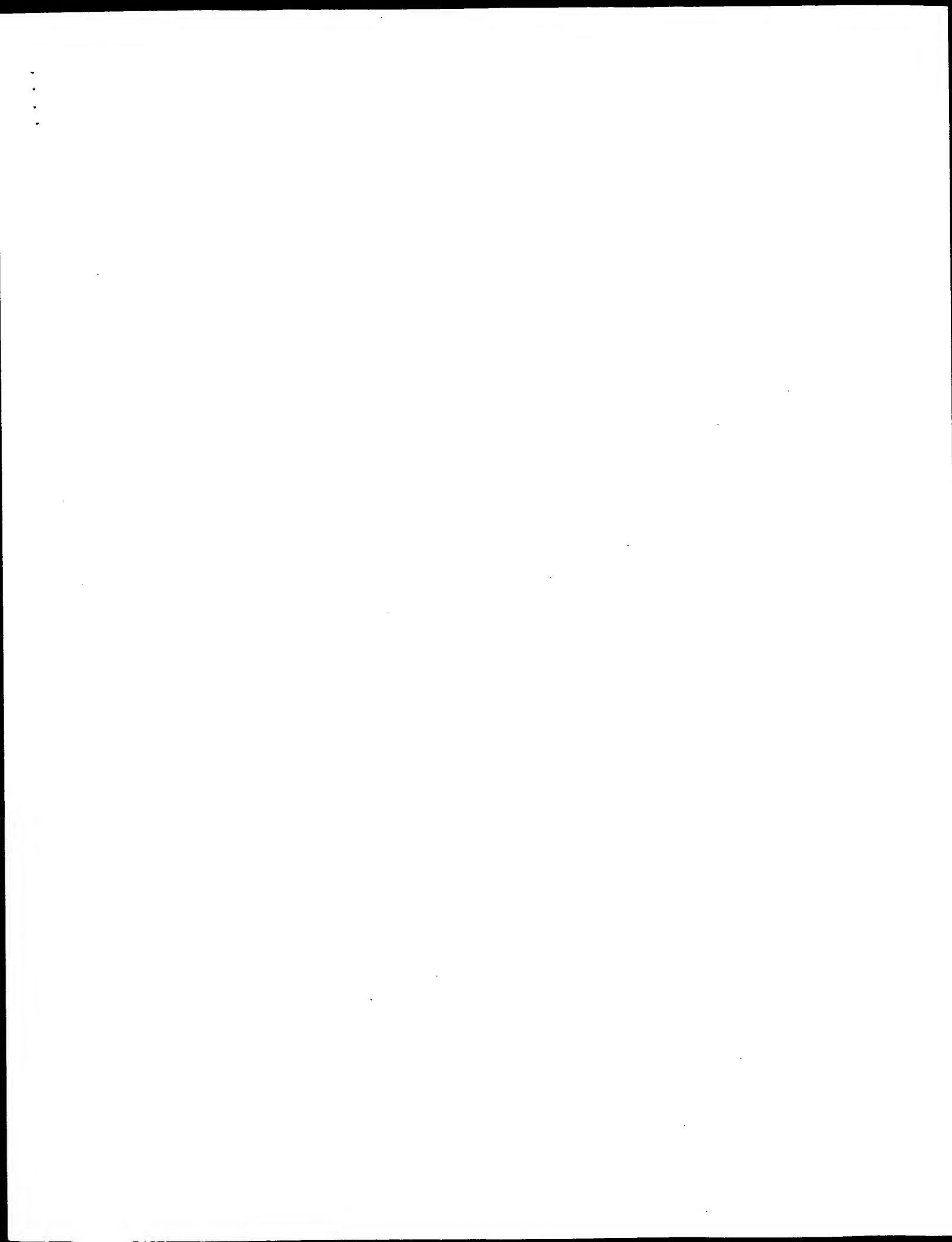
1488 TTCCTCTGTGTAATTGGCCATCAACACACAGTCCTTTTGAAGACATCCATCAGTACAGG 154:
 QY 119 GlnGlnIleLeuAepThrLysSerCysLeuLysAsnLysThrLysGluAenValAspVal 138
 Db ::::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 1548 GAGCAGATC-----AAGCGGCTGAAAGATTCCAGATGATGTG 1583
 QY 139 ProLeuValIleCysGlyAsnLysGlyAspArgAspPheTyArgGluuValAspGlnAtg 158
 Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 1584 CCAATGGTGCTGTGGCGAACAAAGTGACCTGCCGCCTCAC---ACTGTTGAGTCTCGG 1640
 QY 159 GluIleGluInLeuValcilyAspProGlnArgCysAlaTyPheGluIleSerAla 178
 Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 1641 CAGGCCAGGACCTTGCTCGCAGC-----TATGGCATCCCTACATTGAACAATCAGCC 1694
 QY 179 LysLysAsnSerSerLeuAspGlnMetPheArgAlaLeu-----Phe 192
 Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 1695 AAGACCCGACCAGGTGTGGAGGATGCTCTTCTACACACTTAGTACGTGAGATTCGGCAGCAT 1754
 QY 193 AlaMetAlaLyLeu-----ProSerGluMetSerProAspLeuHisAgLysValSer 210
 Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 1755 AAACCTCGGAAACTGAACCCGCCTGATGAGAGTGCCCTGG-CTGCAT-----GAG 1804
 QY 211 ValGlnTyrCyeAspValLeu 217
 Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 1805 CTCGAAGTGTGTGCTGTCTG 1825

RESULT 14
 US-08-884-866A-10
 ; Sequence 10, Application US/08884866A
 ; GENERAL INFORMATION:
 ; APPLICANT: Chien, Shu
 ; APPLICANT: Shyu, John Y-J
 ; TITLE OF INVENTION: GENE THERAPY IN CORONARY ANGIOPLASTY AND
 ; TITLE OF INVENTION: BYPASS
 ; FILE REFERENCE: UCSd1100-1
 ; CURRENT APPLICATION NUMBER: US/08/884,866A
 ; CURRENT FILING DATE: 1997-06-30
 ; PRIOR APPLICATION NUMBER: 60/030,358
 ; PRIOR FILING DATE: 1996-11-08
 ; NUMBER OF SEQ ID NOS: 25
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 10
 ; LENGTH: 450
 ; TYPE: DNA
 ; ORGANISM: Rattus norvegicus
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (49)...(51)
 ; OTHER INFORMATION: nnn = Any nucleic acid triplet, except for UCA,
 ; OTHER INFORMATION: uuc, ucg, ucu, agc, or agu
 US-08-884-866A-10

Alignment Scores:
 Pred. No.: 1,79e-23 Length: 450
 Score: 266.50 Matches: 61
 Percent Similarity: 58.06% Conservative: 29
 Best Local Similarity: 39.35% Mismatches: 54
 Query Match: 18.38% Indels: 11
 DB: Gaps: 3

US-09-709-103-1F1 (1-282) x US-08-884-866A-10 (1-450)

QY 25 TyrArgMetValIleLeuGlySerSerLysValGlyLysThrAlaIleValSerArgPhe 44
 Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 10 TATAAGCTGTGTGTGGCGCGCGCGGTGTGGCGCAANNNGCGCTCACCATCCAGCTG 69
 QY 45 LeuThrGlyArgPheGluAspAlaTyThrProThrIleGluAspPheHisArgLysPhe 64
 Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 70 ATCCAGAACCATTTTGTGACGACNATHAGCCCCACTATAGAGGATTCCTACCGGNAAGCAG 129
 QY 65 TyrsertileArgGlyGluValTyGlnLeuAspIleLeuAspThrSerGlyAsnHisPro 84
 Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 130 GTGTGTCATTGATGGGAGACGTCCTGTTGGACATCTTGATACCGCGCGCTGGAGGAG 189



GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: December 30, 2002, 14:39:31 ; Search time 2190 Seconds
(without alignments)

2085.446 Million cell updates/sec

Title: US-09-709-103-1f1

Perfect score: 1450

Sequence: 1 MKLAMIKRCPSDELSIP.....REKASAGSQAQKXRCVCIS. 282

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-O/cgmt 1/USPTO.spool/TRANSUS09709103/runat.30122002_143926_20664/app.query.fasta_1.455
-DB=EST -OPMT=fastlap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blonum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-USER=TRANSUS09709103 @CGN_1_1_1716 @runat.30122002_143926_20664 -NCRU=6
-ICPU=3 -NO_XLPHY -NO_MMWP -LANGQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELTEXT=7

Database :

EST: *
1: em_estdb: *
2: em_esthum: *
3: em_estin: *
4: em_estnu: *
5: em_estov: *
6: em_estpl: *
7: em_estio: *
8: em_hlc: *
9: gb_estl: *
10: gb_est2: *
11: gb_hlc: *
12: gb_est3: *
13: gb_est4: *
14: gb_est5: *
15: em_estfun: *
16: em_estom: *
17: gb_gss: *
18: em_gss_hum: *
19: em_gss_inv: *
20: em_gss_pln: *
21: em_gss_vrt: *
22: em_gss_fun: *
23: em_gss_mam: *
24: em_gss_mus: *
25: em_gss_other: *
26: em_gss_pro: *
27: em_gss_rtd: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1210.5	83.5	1035	14	BM919341 AGENCOURT
2	1165	80.3	962	13	BM543472 AGENCOURT
3	1144	78.9	1103	14	BM920514 AGENCOURT
4	1121	77.3	965	9	AL533318
5	1120	77.2	742	10	AW028127
6	1114.5	76.9	958	14	BQ719566
7	1100.5	75.9	1053	14	BM921737
8	1064	72.4	649	12	BG085090
9	1038	71.6	1023	13	BM543630
10	1033	71.2	1032	14	BQ067637
11	991.5	68.4	726	12	BF613135
12	987	68.1	655	12	BG711792
13	926	63.9	758	13	BI553776
14	918.5	63.3	1137	14	BM921656
15	917	63.2	814	13	BI596688
16	896.5	61.8	1160	14	BM805574
17	892	61.5	644	13	BG969048
18	890.5	61.4	831	13	BI596637
19	885.5	61.1	710	13	BJ526038
20	884	61.0	1006	14	BQ073742
21	889	59.9	506	13	BM311047
22	867	59.8	1300	11	AK015898
23	862	59.4	947	14	BQ954076
24	859	59.2	648	10	BB636889
25	859	59.2	689	10	BB632699
26	854	58.9	904	14	BQ947936
27	850	58.6	617	13	BI393669
28	832.5	57.4	640	13	BQ706012
29	825	56.9	699	12	BG709229
30	812	56.0	700	13	BI596509
31	812	56.0	700	13	BI596509
32	809	55.8	589	13	BM426066
33	807.5	55.7	888	13	BI754083
34	785.5	54.2	1332	13	BM460899
35	767	52.9	674	13	BI601563
36	748	51.6	675	13	BI545172
37	745	51.4	666	13	BI549939
38	741	51.1	571	13	BI682922
39	732.5	50.5	456	12	BF555822
40	730	50.3	509	10	AW915326
41	703.5	48.5	505	13	BI391213
42	685	47.2	542	9	AA790463
43	674.5	46.5	1438	14	BM807669
44	660.5	45.6	435	10	AW990252
45	655.5	45.2	461	13	BU489219

ALIGNMENTS

RESULT 1
BM919341
LOCUS BM919341 1035 bp mRNA linear EST 12-MAR-2002
DEFINITION AGENCOURT_6715681 NIH_MGC_120 Homo sapiens CDNA clone IMAGE:5748585
5', mRNA sequence.
ACCESSION BM919341
VERSION BM919341.1 GI:19369720
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 1035)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL
COMMENT

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12719 row: j column: 10
High quality sequence stop: 658.

FEATURES

source

1..1035
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5748585"
/clone_lib="NIH MGC_120"
/lab_host="DH10B"
/note="Organ: pooled pancreas and spleen; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source anonymous pool of spleen and pancreas from 28 yo male.. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.5 kb, insert size range 1-2.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 025. Note: this is a NIH_MGC Library."
BASE COUNT 209 a 386 c 293 g 146 t 1 others
ORIGIN

Alignment Scores:

Pred. No.: 1,27e-141 Length: 1035
Score: 1210.50 Matches: 244
Percent Similarity: 95.02% Conservative: 4
Best Local Similarity: 93.49% Mismatches: 8
Query Match: 83.48% Indels: 5
DB: 14 Gaps: 2

US-09-709-103-1F1 (1-282) x BM919341 (1-1035)

QY 1 MetLysLeuAlaAlaMetIleLysLysMetCysProSerAspSerGluLeuSerIlePro 20
Db 187 ATGAACTGGCCGCGATGATCAAGAGATGTGCCGAGCGACTCGAGCTGAGTATCCCG 246
QY 21 AlaLysAsnCysTyrArgMetValIleLeuGlySerLysValGlyLysThrAlaIle 40
Db 247 GCCAAGAACTGCTATCGCATGTCTCTCGCTCGTCCAGGTGGCGCAAGCGGCATC 306
QY 41 ValSerArgPheLeuThrGlyArgPheGluAspAlaTyrThrProThrIleGluAspPhe 60
Db 307 GTGTGCGCGCTTCTCCACCGCGCTTCGAGGACGCTTACAGCGCTACCATCGAGGACTTC 366
QY 61 HisArgLysPheTyrSerIleArgGlyGluValTyrGlnLeuAspIleLeuAspThrSer 80
Db 367 CACCGCAAGTTCTACTCCATCCCGCGGAGGTTCACAGCTCGACATCTCGACAGCTCC 426
QY 81 GlyAsnHisProPheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIle 100
Db 427 GGCAACACCGCTTCCCGCGCATCGCGCGCTCTCCATCTCCACAGAGAGCTTTTCATC 486
QY 101 LeuValPheSerLeuAspAsnArgAspSerPheGluGluValGlnArgIleArgGlnGln 120
Db 487 CTGTGTTCAGTCTGGACAAACCGCGACTCTCTCGAGAGGTGCGCGGTTCAGGAGCAG 546
QY 121 IleLeuAspThrLysSerCysLeuLysAsnLysThrLysGluAsnValAspValProLeu 140
Db 547 ATCTCGACACCAAGTCTTCCTCCAGACAAACACAGAGAACGTGACGTGCCCTG 606
QY 141 ValIleCysGlyAsnLysGlyAspArgAspPheTyrArgGluValAspGlnArgGluIle 160
Db 607 GTCATCTGCGGCAACAGGGGTGACCGCGACTTCTACCGGAGGTGTGACCGCGGAGATC 666

QY 161 GluGlnLeuValGlyAspAspProGlnArgCysAlaTyrPheGluIleSerAlaLysLys 180
Db 667 GAGCAGCTGGTGGCGAGAGACCCCGAGCGCTGGCGCTACTTCGAGATCTCGCCCAAGAG 726
QY 181 AsnSerSerLeuAspGlnMetPheArgAlaLeuPheAlaMetAlaLysLeuProSerGlu 200
Db 727 AACAGCAGCTCGACACAGATGTTCGCGCGCTCTTCGGCATGTGGCAAGTGTCCCGAG 786
QY 201 MetSerProAspLeuHisArgLysValSerValGlnTyrCysAspValLeuHisLysLys 220
Db 787 ATGAGCCAGACCTCCACCGCAAGGTCTCGGTGAGTACTCGGACGTGTCACAAAGAA 846
QY 221 AlaLeuArgAsnLysLysLeuLeuArgAlaGlySerGly-GlyGly-GlyGlyAspProG 240
Db 847 GCGCTGCGGAAACAGAAACTGCTGCGGCGGCAACGCGCGCGCGCGCGCGCGCGCGCG 906
QY 240 LysAspAlaPheGlyIleValAla-----ProPheAlaArg---ArgProSerValHis 256
Db 907 GCGAGCGCCCTTTGGCCATTCGTGGCAACCCCTTTTCGCGCGCGCGCGCGCGCGCGCG 965
RESULT 2
BM543472 962 bp mRNA linear EST 20-FEB-2002
LOCUS AGENCOURT 6492614 NIH_MGC_124 Homo sapiens cDNA clone IMAGE:5726554
DEFINITION 5', mRNA sequence.
ACCESSION BM543472
VERSION BM543472.1 GI:18773895
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 962)
NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Invitrogen
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12719 row: d column: 11
High quality sequence start: 18
High quality sequence stop: 692.
FEATURES
Location/Qualifiers
source
1..962
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5726554"
/clone_lib="NIH MGC_124"
/tissue_type="hippocampus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMV-SPORT6; Site 1: EcoRV (destroyed); Site 2: NotI; RNA source male hippocampus, age 27. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 0.9-4 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 012."
BASE COUNT 185 a 363 c 272 g 141 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 6.21e-136 Length: 962
Score: 1165.00 Matches: 234
Percent Similarity: 92.55% Conservative: 2
Best Local Similarity: 91.76% Mismatches: 9
Query Match: 80.34% Indels: 10

DB: 13 Gaps: 1

US-09-709-103-1F1 (1-282) x BM543472 (1-962)

QY 1 MetLysLeuAlaMetIleLysLysMetCysProSerAspSerGluLeuSerIlePro 20

DB 224 ATGAACTGACCGCGATGATCAAGAAAGATGCGCCAGCTCGAGTGTATCCCG 283

QY 21 AAlaLysAsnCysTyrArgMetValIleLeuGlySerSerLysValGlyThrAlaIle 40

DB 284 GCCAAGAACTGCTATGCGATGCGATCTCGGTCTCCAGGTGGCAAGCGCCATC 343

QY 41 ValSerArgPheLeuThrGlyArgPheGluAspAlaTyrThrProThrIleGluAspPhe 60

DB 344 GTGTCCGCTTCTCTCAACCGCGCTTCGAGAGCGCTTACACCTCCTACCTCAGAGACTTC 403

QY 61 HisArgLysPheThrSerIleArgGlyGluValTyrGlnLeuAspIleLeuAspThrSer 80

DB 404 CACCGCAAGTCTACTCCATCCGCGGAGGTCTTACCACTCGACATCCTCGACACGCTC 463

QY 81 GlyAsnHisProPheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIle 100

DB 464 GCGAACCACCGCTTCCCGCCATGCGCGCTTCTCATCTTCACAGAGACGTTTTCATC 523

QY 101 LeuValPheSerLeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGlnGln 120

DB 524 CTGGTGTTCAGTCTGACAAACCGCGACTCTTCGAGAGGTGAGCGGCTCAGGACGACG 583

QY 121 IleLeuAspThrLysSerCysLeuLysAsnLysThrLysGlnAsnValAspValProLeu 140

DB 584 ATCTCTGACACCAAGTCTTCTGCTTCAGAAACCAAAACCAAGAGACGTGACGTCCTCG 643

QY 141 ValIleCysGlyAsnLysGlyAspArgAspPheTyrArgGluValAspGlnArgGluIle 160

DB 644 GTCTATCTGCGGCAACAAGGTGACCGCGACTTCTACCGGAGGTGACCAAGCGCGATC 703

QY 161 GluGlnLeuValGlyAspAspProGlnArgCysAlaTyrPheGluIleSerAlaLysLys 180

DB 704 GACGACCTGGTGGCGACGACCCCGCGCTGCGCTTCTGAGATCTCGGCGCAAGG 763

QY 181 AsnSerSerLeuAspGlnMetPheArgAlaLeuPheAlaMetAlaLysLeuProSerGlu 200

DB 764 AACAGAGCCTTGACCAAGATGTCGCGCGCTTCTGCGACATGCGCCAGCTGCCACGCG 823

QY 201 MetSerProAspLeuHisArgLysValSerValGlnTyrCysAspValLeuHisLysLys 220

DB 824 ATGAGCCCAAGCTTCCAGCGCAAGTCTCGGTCACTGCGACGTCTGCAAGAA 883

QY 221 AAlaLeuArgAsnLysLysLeuLeuArgAlaGlySerGlyGlyGlyGlyAspProG 240

DB 884 GCGCTGGGGGACAAAGAACTGCTGGGGGGCGGGGCAACCGCGCGCGCGCGCGCA 943

QY 240 LysAspAlaPheGlyIleValAlaProPheAlaArgArgPro 253

DB 944 ACCCG-----GGGCGACGCGCT 960

RESULT 3

LOCUS BM920514 1103 bp mRNA linear EST 12-MAR-2002

DEFINITION AGENCOURT 6709473 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5750507

ACCESSION BM920514

VERSION BM920514.1 GI:19370893

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1 (bases 1 to 1103)

AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgabs-r@mail.nih.gov

Tissue Procurement: Life Technologies, Inc.

CDNA Library Preparation: Life Technologies, Inc.

DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Plate: LLAM2781 row: j column: 12

High quality sequence stop: 676.

Location/Qualifiers

1..1103

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:5750507"

/clone_1ib="NIH_MGC_122"

/lab_host="DH10B"

/note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of 24 week female lung, 16 week female spleen, and 20-22 week male spleens. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 026. Note: this is a NIH MGC Library."

BASE COUNT 230 a 402 c 314 g 156 t 1 others

ORIGIN

Alignment Scores:

Pred. No.: 3,42e-133 Length: 1103

Score: 1144.00 Matches: 246

Percent Similarity: 88.69% Conservative: 5

Best Local Similarity: 86.93% Mismatches: 24

Query Match: 78.90% Indels: 9

DB: 14 Gaps: 3

US-09-709-103-1F1 (1-282) x BM920514 (1-1103)

QY 1 MetLysLeuAlaMetIleLysLysMetCysProSerAspSerGluLeuSerIlePro 20

DB 201 ATGAACTGACCGCGATGATCAAGAAAGATGCGCCAGCTCGAGTGTATCCCG 260

QY 21 AAlaLysAsnCysTyrArgMetValIleLeuGlySerSerLysValGlyThrAlaIle 40

DB 261 GCCAAGAACTGCTATGCGATGCGATCTCGGTCTCCAGAGGTGGCAAGCGCCATC 320

QY 41 ValSerArgPheLeuThrGlyArgPheGluAspAlaTyrThrProThrIleGluAspPhe 60

DB 321 GTGTCCGCTTCTCTCAACCGCGCTTCGAGAGCGCTTACACGCTTACCTCAGAGACTTC 380

QY 61 HisArgLysPheThrSerIleArgGlyGluValTyrGlnLeuAspIleLeuAspThrSer 80

DB 381 CACCGCAAGTCTACTCCATCCGCGGAGGTCTTACCAAGCTCGACATCTCGACACGCTC 440

QY 81 GlyAsnHisProPheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIle 100

DB 441 GCCAACCACCGCTTCCCGCCATGCGCGCTTCTCATCTTCACAGAGACGTTTTCATC 500

QY 101 LeuValPheSerLeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGlnGln 120

DB 501 CTGGTGTTCAGTCTGACAAACCGGTGACTCTTCAGAGAGGTGACGCGCTCAGGCAACG 560

QY 121 IleLeuAspThrLysSerCysLeuLysAsnLysThrLysGlnAsnValAspValProLeu 140

DB 561 ATCTCTGACACCAAGTCTTCTGCTTCAGAAACCAAGAGAGACGTGACGTCCTCG 620

QY 141 ValIleCysGlyAsnLysGlyAspArgAspPheTyrArgGluValAspGlnArgGluIle 160

DB 621 GTCTATCTGCGGCAACAAGGTGACCGCGACTTCTTACCGGAGGTGACCAAGCGCGATC 680

QY 161 GluGlnLeuValGlyAspAspProGlnArgCysAlaTyrPheGluIleSerAlaLysLys 180

```

Db 681 GAGCAGCTGGTGGGACGACCCCGAGCGCTGCGCTACTTCGAGATCTCGGCCAAGAG 740
QY 181 AenSerSerLeuAspGlnMetPheArgAlaLeuPheAlaMetAlaLysLeu-ProSerG 200
Db 741 AA-AGCAGCCTGGACAGATGTTCCGCGCTCTTCGCGCATGGGCGAGCTCCCGGCGGA 799
QY 200 uMetSerProAspLeuHisArgLysValSerValGlnTyrCysAspValLeuHis-LysL 220
Db 800 GATGAGCCGAGACCTTCGACCGCAAGTCTCGGTGAGTACTGCGACGTCGTCACAAAAA 859
QY 220 ysAlaLeuArgAsnLys-LysLeuLeuArgAlaGlySerGlyGlyGlyGlyAspPro 239
Db 860 AGGCTTCGCGAACAAGAGCTGTTCCGGGCGCGGCAACCGCGCGGCGCGGCCCA 919
QY 240 GlyAspAla-----PheGlyLeValAlaProPhe---AlaArgArgProSerValHis 256
Db 920 ACCGGGGCGCAACCCCTTTGGCATCTGGGAGCCCTTTCCCGCGCGGGGCCCAAGGTACAC 979
QY 257 -SerAspLeuMetTyrIleArgGluLysAlaSer---AlaGlySerGlnAlaLysAspLy 275
Db 980 AGGGAACCTTATGAAATCCCCCCCCAAAGGCAACGCCCGGNCAGCCAGGCGCAAGAGAAA 1039
QY 275 s 275
Db 1040 A 1040

RESULT 4
AL533318 965 bp mRNA linear EST 13-FEB-2001
LOCUS AL533318 LTI_FL015_Brn1 Homo sapiens cDNA clone CS0DN003YJ19 5
DEFINITION prime mRNA sequence.
ACCESSION AL533318
VERSION AL533318.1 GI:12796811
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Li, W.B., Gruber, C., Jesse, J., and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers
1. .965
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DN003YJ19"
/clone_lib="LTI_FL015_Brn1"
/sex="male"
/tissue_type="Adult brain"
/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was constructed by Life Technologies.
Contact : Feng Liang Life Technologies, a division of
Invitrogen 9800 Medical Center Drive Rockville, Maryland
20850, USA Fax : (1) 301 610 8371 Email :
fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT 189 a 352 c 273 g 144 t 7 others
ORIGIN

Alignment Scores:
Pred. No.: 2.21e-130 Length: 965
Score: 1121.00 Matches: 237
Percent Similarity: 95.22% Conservative: 2
Best Local Similarity: 94.42% Mismatches: 11

```

```

Query Match: 77.31% Indels: 5
DB: 9 Gaps: 1
US-09-709-103-1f1 (1-282) x ALS33318 (1-965)
QY 1 MetLysLeuAlaAlaMetIleLysLysMetCysProSerAspSerGluLeuSerIlePro 20
Db 191 ATGAAACTGGCGCGGATGATCAAGAAGATGTGCCGAGGAGCTCGAGCTGAGTATCCCG 250
QY 21 AlaLysAsnCyfTzArgMetValIleLeuGlySerSerLysValGlyLysThrAlaIle 40
Db 251 GCCAAGAACTGCTATCGCATGTCTCTCGCTCGCTCGTCCAAAGTGGCAAGACGCCATC 310
QY 41 ValSerArgPheLeuThrGlyArgPheGluAspAlaTyrThrProThrIleGluAspPhe 60
Db 311 GTGTCCGCTTCTTCCACCGCGCGCTTCGAGGACGCTTACACGCTTACCATCCATCAGGACTTC 370
QY 61 HisArgLysPheTyrSerIleArgGlyGluValTyrGlnLeuAspIleLeuAspThrSer 80
Db 371 CACCGCAAGTTCTACTCCATCCGCGGAGGTCTACAGCTCGACATCTTCGACACGTC 430
QY 81 GlyAsnHisProPheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIle 100
Db 431 GGCAACACCCGTTCCCGCCCATGCG-CGCTCTCTCATCTCTCACAGGAGACGTTTTCATC 489
QY 101 LeuValPheSerLeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGln 120
Db 490 CTGGTGTTCAGTCTGGACAACCGCGACTCTTTCGAGGAGGTGCAGCGGCTCAGGCGAG 549
QY 121 IleLeuAspThrLysSerCysLeuLysAsnLysThrLysGluAsnValAspValProLeu 140
Db 550 ATCTCGACACCAAGTCTTTCCTCAAGAACAAACCAAGGAGAACGTGGACGTGCCCTG 609
QY 141 ValIleCysGlyAsnLysGlyAspArgAspPheTyrArgGluValAspGlnArgGluIle 160
Db 610 GTCATCTCGCGCAACAAGGTGACCGCGACTTCTACCGGAGGTGCAGCGGCTCAGGCGAG 669
QY 161 GluGlnLeuValGlyAspAspProGlnArgCysAlaTyrPheGluIleSerAlaLysLys 180
Db 670 GAGCAGCTGTGGCGGACGACACCCCGAGCTCTTCGAGATCTTCGAGATCTCGGCCAAGA 729
QY 181 AsnSerSerLeuAspGlnMetPheArgAlaLeuPheAlaMetAlaLysLeuProSerGlu 200
Db 730 AAMAGCAGCTGACACAGATGTTCCGCGCTCTTCGCGCATGGCCCAAGCTGCCAGCGAG 789
QY 201 MetSerProAspLeuHisArgLysValSerValGlnTyrCysAspValLeuHisLysLys 220
Db 790 ATGAGCCACAGACTGCACCGCAAGTC-TCGGTGCAGTACTGCGACGTCTGCAMAARAAR 848
QY 221 AlaLeuArgAsnLysLysLeuLeuArgAlaGlySerGlyGlyGlyGlyAspProGly 240
Db 849 GCGCTCGCGAACA--AARCTGTGCGGCGGCGAGCGS--GGCGGCGGCGGCGGCGCG 905
QY 241 AspAlaPheGlyLeuValAlaProPheAlaArg 251
Db 906 ---AGCTTTGCTGTCGACCTTTCGCGCGG 935

RESULT 5
AL533318 742 bp mRNA linear EST 27-OCT-1999
LOCUS AL533318 NCI CGAP Kid11 Homo sapiens cDNA clone IMAGE.2530668 3'
DEFINITION similar to TR:O55626 O35626 RAS, DEXAMETHASONE-INDUCED 1 ;, mRNA
sequence.
ACCESSION AL533318
VERSION AL533318.1 GI:5886883
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 742)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

```

JOURNAL
COMMENT

Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaabs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrip/image/image.html
Seq primer: -400P from Gibco
High quality sequence stop: 445.
Location/Qualifiers

FEATURES

source

1..742
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2530668"
/clone_1db="NCI CGAP_Kid11"
/lab_host="DH10B"
/note="Organ: Kidney; Vector: pT73D-Pac (Pharmacia) with
a modified polylinker; Site 1: Not I; Site 2: Eco RI;
Plasmid DNA from the normalized library NCI_CGAP_Kid3 was
prepared, and ss circles were made in vitro. Following HAP
purification, this DNA was used as tracer in a subtractive
hybridization reaction. The driver was PCR-amplified cDNAs
from a pool of 5,000 clones made from the same library
(cloneids 1323376-1323911, 1456007-1456775, and
1500552-1502855). Subtraction by Bento Soares and M.
Fatima Bonaldo. "
BASE COUNT 149 a 260 c 208 g 122 t 3 others
ORIGIN

Alignment Scores:

Pred. No.: 1.97e-130 Length: 742
Score: 1120.00 Matches: 220
Percent Similarity: 95.74% Conservative: 5
Best Local Similarity: 93.62% Mismatches: 10
Query Match: 77.24% Indels: 0
DB: Gaps: 0

US-09-709-103-1F1 (1-282) x AM028127 (1-742)

QY 1 MetLysLeuAlaAlaMeCilIeLysLysMetCysProSerAspSerGluLeuSerIlePro 20
DB 35 ATGAACACTGCGCCGATGATCAAGAGATGCGCCGAGGAGCTCGAGCTGATGCCG 94
QY 21 AlAluysAsnCySTyArGMeValIleuGlySerSerIyValGlyLysThrAlaIle 40
DB 95 GCCAAGAACTGCTATCGCATGTGTCATCCTCGCTCCAGAGTGGGCAAGCGGCCATC 154
QY 41 ValSerArGPhLeuThrGlyArGPhGluAspAlaIyTrhProThrIleGluAspPhe 60
DB 155 GGTGCGCGCTTCTCACTCGCGCTTCGAGGAGCGCTACACCGCTACCATCATGAGGACTTC 214
QY 61 HtsArGlySPheTySerIleArGlyGlyValIyTrGlnLeuAspIleLeuAspPhsSer 80
DB 215 CACCGCAAGTTTACTTCATCCGCGCGAGGCTACAGCTCGACATCTCGACACGCTC 274
QY 81 GlYAsnHisProPheProAlaMetArGArGleuSerIleLeuThGlyAspValPheIle 100
DB 275 GGCAACACACCGCTTCCCGCCGATGGCGGCGCTCTCCATCTTCACAGAGAGCTTTTATC 334
QY 101 LeuValPheSerLeuAspAsnArGAspSerPheGluGluValGlnArGLeuArgGln 120
DB 335 CTGGGTTCAGTCTGAGAACCGGACGCTCTTCGAGGAGGTGCAAGCGGCTCAGGACGAG 394
QY 121 IleLeuAspThrLysSerCysLeuLysAsnLysThrLysGluAspValAspValProIle 140
DB 395 ATCTCTGACACCAAGTCTTGCTTCAAGAACCAACCAAGAGAGAGTGGAGCGTCCCTG 454
QY 141 ValIleCySGlyAsnLysGlyAspArGAspPhetyrArGlyValAlaSpGlnArgGluIle 160

DB 455 GTCATCTGCCGCAACAAAGGCTGACCCGACTTTCACCGGAGGTGGACCAAGCGCGAGATC 514

QY 161 GluGlnLeuValGlyAspAspProGlnArGySAIATyPheGluIleSerAlaLysLys 180

DB 515 GAGCAGCTGTGGGCGAGCGACCCCGAGCGCTGGCTTCTGCAATCTGCAAGTCTGCGCAAGAG 574

QY 181 AsnSerSerLeuAspGlnMetPheArGAlaLeuPheAlaMetAlaLysLeuProSerGlu 200

DB 575 AACGACGAGCTTGACCGAGAGTTCGCGGGCTTTTGGCATGGCCAAAGTGGCCACGAG 634

QY 201 MetSerProAspLeuHisArGlyValSerValGlnTyrCysAspValLeuHisLysLys 220

DB 635 ATGAGCCCAACCTGCAACCGCAGGCTCTGNGTGCAGTACCTGGACGCTGCTGCACAGAAG 694

QY 221 AlaleuArGAsnLysLysLeuArGAlaGlySerGlyGly 235

DB 695 GCGCTGCCGAACAGAGAGCTGTGCGTCCGACGACGCGGCTGCT 739

RESULT 6 958 bp mRNA linear EST 16-JUL-2002

LOCUS BQ719566

DEFINITION AGENCOUNT 8305191 lupski_sympathetic_trunk Homo sapiens CDNA clone

IMAGE:6193170 5', mRNA sequence.

ACCESSION BQ719566

VERSION BQ719566.1 GI:21858463

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1 (bases 1 to 958)

NITH-MGC http://mgc.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

CONTACT: Robert Strausberg, Ph.D.

Email: cgaabs-r@mail.nih.gov

Tissue Procurement: Dr. James R. Lupski

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM13596 row: f column: 19

High quality sequence stop: 514.

FEATURES

source

1..958

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:6193170"

/clone_1db="Lupski_sympathetic_trunk"

/sex="male"

/tissue_type="sympathetic trunk"

/dev stage="adult, 16 yr"

/lab_host="DH10B"

/note="Vector: pCMV-SPORE (Life Technologies); Site 1:

NotI; Site 2: SalI; cDNA made by oligo-dT priming.

Directionally cloned using the following adaptors:

5'-TCGACGACGAGCGCTCG-3' and

5'-GACATGCTTAGATCGGACGCGCGCTT(15'-3'. Size selected >

1 kb for average insert length 1.9 kb. This is a primary

library, non-amplified. Library constructed by Life

Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor

College of Medicine); available through Life

Technologies."

BASE COUNT 185 a 339 c 275 g 159 t

ORIGIN

Alignment Scores:

Pred. No.: 1.44e-129 Length: 958

Score: 1114.50 Matches: 226

Percent Similarity: 94.61% Conservative: 2

Best Local Similarity: 93.78% Mismatches: 10					
Query Match: 76.86%					
DB: Indels: 3					
Gaps: 1					
US-09-709-103-1F1 (1-282) x BQ719566 (1-958)					
Oy	15	SerGluleuSerleProAlalysAsnCysTyrArgMetValIleLeuGlySerSerlys	34		
Dd	1	TCCGAGCTGATATCCCGCAAGAACTGTATCGCATGTCATCCTCGCTCGTCCAAG	60		
Oy	35	VaiGlyLysThrAlalleValSerArgPheLeuThrGlyArgPheGluAspAlaTyrThr	54		
Dd	61	GTTGGGCAAGACGGCCATCGTGTCGCCTTCTCACCGCGCGCTTCGAGACGCCCTACACG	120		
Oy	55	ProThrileGluAspPheHisArgLysPheTyrSerIleArgGlyGluValTyrGlnLeu	74		
Dd	121	CCTACCATTGAGACTTCCACCCNAGTTCATCTCATCCGCGGCGCTTACAGATC	180		
Oy	75	AspIleLeuAspThrSerGlyAsnHisProPheProAlaMetArgArgLeuSerIleLeu	94		
Dd	181	GACATCTCGACAGCTCCGCAACACCCGCTTCCCGCCATCGCGCGCTTCCATCCTC	240		
Oy	95	ThrGlyAspValPheIleLeuValPheSerLeuAsnArgAspSerPheGluGluVal	114		
Dd	241	ACAGGAGACGCTTTTCATCTCTGTTGTCAGTCTCGAACAACCGGACCTTCGAGAGGTG	300		
Oy	115	GlnArgLeuArgGlnGlnIleLeuAspThrLysSerCysLeuLysAsnLysThrLysGlu	134		
Dd	301	CAGCGCTCAGCGNCGAGATCTCGACACCAAGTCTTGCTCAAGAACAACACCAAGNG	360		
Oy	135	AsnValAspValProLeuValIleCysGlyAsnLysGlyAspArgAspPheTyrArgGlu	154		
Dd	361	AACGTGACGTGCCCCCTGGTCACTCGCGCAACAAGGGTGACCGGACTTCTACCGGAG	420		
Oy	155	ValAspGlnArgGluIleGluGlnLeuValGlyAspAspProGlnArgCysAlaTyrPhe	174		
Dd	421	GTGCACGACGCGAGATCGAGCAGCTGTGTGGCGGACGACCCCAGCGCTCGCCCTACTTC	480		
Oy	175	GluIleSerAlalysLysAsnSerSerLeuAspGlnMetPheArgAlaLeuPheAlaMet	194		
Dd	481	GAGATCTCGGCCAAGAAGAACAGCAGCTTGACACAGATGTTCCGCGCGCTCTTCGCCATG	540		
Oy	195	AlalysLeuProSerGluMetSerProAspLeuHisArgLysValSerValGlnTyrCys	214		
Dd	541	GCCAACTGCCAGCAGATGACCCAGACCTGCCCGCAAGGTCTCGGTGAGTAATGC	600		
Oy	215	AspValLeuHisLysLysAlaLeuArgAsnLysLysLeuLeuArg-AlaGlySerGly-G	234		
Dd	601	GACGTGTGCACAAGAGCGCTGCGGAACAAGAAAGTGTGCGGGCGCGCAACGGAAG	660		
Oy	234	IyGlyGlyGlyAspPro--GlyAspAlaPheGlyIleValAlaProPheAlaArgAtq	252		
Dd	661	GCCGCGCGCTGACCCCGGGGAAGCTTTTGGGATCGGGGCGCCCTTTCGCGCGCG	719		
RESULT 7					
BM921737					
LOCUS					
DEFINITION					
AGENCOURT_6708101 NIH_MGC_115 Homo sapiens cdna clone IMAGE:5753301					
5', mRNA sequence.					
ACCESSION					
VERSION					
BM921737					
KEYWORDS					
EST.					
SOURCE					
human.					
ORGANISM					
Homo sapiens					
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
REFERENCE					
1 (bases 1 to 1053)					
AUTHORS					
NIH-MGC http://mgi.nci.nih.gov/					
TITLE					
National Institutes of Health, Mammalian Gene Collection (MGC)					
JOURNAL					
Unpublished (1999)					
COMMENT					
Contact: Robert Strausberg, Ph.D.					
Email: cgabs-i@mail.nih.gov					
Tissue Procurement: Life Technologies, Inc					
CDNA Library Preparation: Life Technologies, Inc.					
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)					
DNA Sequencing by: Agencourt Bioscience Corporation					
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:					
http://image.llnl.gov					
Plate: LLAM12788 row: n column: 22					
High quality sequence stop: 600.					
FEATURES					
Location/Qualifiers					
1..1053					
source					
/organism="Homo sapiens"					
/db_xref="taxon:9606"					
/clone="IMAGE:5753301"					
/clone_lib="NIH_MGC_115"					
/lab_host="DHIOB"					
/note="Organ: pooled brain, lung, testis; Vector: pCMV-SPORT6; site: 1: NotI; Site 2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27; 1 male lung, age 27; and 1 male testis, age 69. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.8 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 021. Note: this is a NIH MGC Library."					
BASE COUNT					
215 a 390 c 297 g 150 t					
ORIGIN					
Alignment Scores:					
Pred. No.: 9,75e-128 Length: 1053					
Score: 1100.50 Matches: 228					
Percent Similarity: 88.26% Conservative: 5					
Best Local Similarity: 86.36% Mismatches: 22					
Query Match: 75.90% Indels: 9					
DB: 14 Gaps: 3					
US-09-709-103-1F1 (1-282) x BM921737 (1-1053)					
Oy	1	MetLysLeuAlaAlaMetIleLysLysMetCysProSerAspSerGluLeuSerIlePro	20		
Dd	200	ATGAACTGGCCCGGATGATCAAGAAGATGTGCCGAGCATCTCGAGCTGAGTATCCG	259		
Oy	21	AlalysAsnCysTyrArgMetValIleLeuGlySerSerLysValGlyLysThrAlaIle	40		
Dd	260	GCCAAAGTGTATCGCATGCTCGCTCGCTCGTCAAGGTGGCAAGCGCCATC	319		
Oy	41	ValSerArgPheLeuThrGlyArgPheGluAspAlaTyrThrProThrIleGluAspPhe	60		
Dd	320	GTGTGCGGCTTCTCCACCGCGCTTCGAGAGCGCTACACGCCCTACCATCGAGGACTTC	379		
Oy	61	HisArgLysPheTyrSerIleArgGlyGluValTyrGlnLeuAspIleLeuAspThrSer	80		
Dd	380	CACCGCAAGTCTTACTTCATCCGCGCGAGGCTTACCACTCGATCTCGACAGCTCC	439		
Oy	81	GlyAsnHisProPheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIle	100		
Dd	440	GGCAACACCGCTTCCCGCCATCGCGGCTCTCCATCTCCACAGGACGCTTTTCATC	499		
Oy	101	LeuValPheSerLeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGlnGln	120		
Dd	500	CTGTGTTCAGTCTGGAACAACCGGACTCTCTTCGAGAGGTGACCGGCTCGAGCAGCAG	559		
Oy	121	IleLeuAspThrLysSerCysLeuLysAsnLysThrLysGluAsnValAspValProLeu	140		
Dd	560	ATCTCTGACACCAAGTCTTTCCTCAAGAACAAAC			

QY 181 AsnSerSer-LeuAspGlnMetPheArgAlaLeuPheAlaMetAlaIleLeuProSerG1 200
 Db 740 AACAGCAGCCCTGACCAAGATGTTCCGGCGGCTCTCCGCGCATGCCAACCTGCCAGCGA 799
 QY 200 uMetSerProAspLeuHisArgIys--ValSerValGlnTyrCysAspValLeuHisIly 219
 Db 800 GATGAGCCCAAGACCTGACCGCAAGGATCTCGGGGCAAGTACTGCCGACCTGTCGACAA 859
 QY 219 sIys---AlaLeuArgAsnIlyLeuLeuArg-AlaGlySer-----G 233
 Db 860 AAAAGGGGCTGCGGGAACAAAAAACTGCTGCGGGCCGTAACCCGTTGGCGGCGCGG 919
 QY 233 IyGlyIyGlyIyGlyAspProGlyAspAlaPheGlyIleValAlaProPheAlaArgArgp 253
 Db 920 GAGGAACCGGGGGAACCCCTTTGGGGAATTCGGGGGCAACCTTGCTGGGGCGCGGAGC 979
 QY 253 roSerVal 255
 Db 980 CAAGGTA 987
 RESULT 8
 LOCUS BG085090 649 bp mRNA linear EST 26-JAN-2001
 DEFINITION H3108E05-5 NIA Mouse 15K cDNA Clone Set Mus musculus cDNA clone
 H3108E05 5', mRNA sequence.
 ACCESSION BG085090
 VERSION BG085090.1 GI:12567654
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 649)
 Kargul,G.J., Dudekula,D.B., Qian,Y., Lim,M.K., Jaraadat,S.A., Tanaka
 T.S., Carter,M.G. and Ko,M.S.H.
 Verification and initial annotation of NIA mouse 15K cDNA clone set
 Unpublished (2001)
 Other ESTs: H3108E05-3
 Contact: George J. Kargul
 Laboratory of Genetics
 National Institute on Aging/National Institutes of Health
 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
 Email: cdna@lgsun.grc.nia.nih.gov
 This clone set has been freely distributed to the community. Please
 visit http://lgsun.grc.nia.nih.gov/cDNA/15K.html for details.
 Plate: H3108 row: E column: 05
 Seq primer: -21M13 Reverse
 High quality sequence stop: 649
 POLYA=No.

FEATURES
 source

1. 649 Location/Qualifiers
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:H3108E05-5"
 /db_xref="taxon:10090"
 /clone="H3108E05"
 /clone_11b="NIA Mouse 15K cDNA Clone Set"
 /sex="Clones arrayed from a variety of cDNA libraries"
 /dev_stage="Clones arrayed from a variety of cDNA libraries"
 /lab_host="DH10B"
 /note="vector: pSPORT1, Site 1: SalI, Site 2: NotI. This clone is among a rearranged set of 15,247 clones from 11 embryo cDNA libraries (including preimplantation stage embryos from unfertilized egg to blastocyst, embryonic part of E7.5 embryos, extraembryonic part of E7.5 embryos, and E12.5 female mesonephros/gonad) and one newborn ovary cDNA library. Average insert size 1.5 kb. All source libraries are cloned unidirectionally with Oligo(dT)-Not primers. References include: (1) Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray, 2000,

Proc. Natl. Acad. Sci. U S A. 97: 9127-9132; (2)
 Large-scale cDNA analysis reveals phased gene expression patterns during preimplantation mouse development, 2000, Development, 127: 1737-1749; (3) Genome-wide mapping of unselected transcripts from extraembryonic tissue of 7.5-day mouse embryos reveals enrichment in the t-complex and under-representation on the X chromosome, 1998, Hum Mol Genet 7: 1967-1978."

BASE COUNT 157 a 182 c 177 g 133 t
 ORIGIN

Alignment Scores:
 Pred. No.: 1.85e-123 Length: 649
 Score: 1064.00 Matches: 212
 Percent Similarity: 98.62% Conservative: 2
 Best Local Similarity: 97.70% Mismatches: 1
 Query Match: 73.38% Indels: 2
 DB: 12 Gaps: 0

US-09-709-103-1F1 (1-282) x BG085090 (1-649)

QY 18 SerIleProAlaIysAsnGlyTyrArgMetValIleuGlySerSerIlyValGlyIys 37
 Db 3 AGTATCCGCGCAAGAACTGTACAGATGTCATCTCGCTCATCAAGTGGCAAG 62
 QY 38 ThrAlaIleValSerArgPheLeuThrGlyArgPheGluAspAlaTyrThrProThrIle 57
 Db 63 ACGGCATTTGTGTGCGCTTCTCAGCGGCCGTTTCAAGATGTTACACCCCTACATC 122
 QY 58 GluAspPheHisArgIysPheTyrSerIleArgGlyGluValTyrGluLeuAspIleu 77
 Db 123 GAGGACTTCCACCAAAAGTTTATTCATCCGCGCGCAAGCTTCCAGTGGACATCTG 182
 QY 78 AspThrSerGlyAsnHisProPheProAlaMetArgArgLeuSerIleuThrGlyAsp 97
 Db 183 GACACATCGGCAATCATCTCCGCGCATGGGCGCTCTTATCTCACAGGAGAC 242
 QY 98 ValPheIleLeuValPheSerLeuAspAsnArgPhePheGluGluValGlnArgLeu 117
 Db 243 GTTTTCATTTCTGGTGTTCAGCTTAGACACCGGCATCTTTCGAAGAGTGCAAGGCTC 302
 QY 118 ArgGlnGlnIleLeuAspThrIlySerCysLeuIysAsnIlyThrIlyGluAsnValAsp 137
 Db 303 AACAGCAGATCTTAGACACCAAGCTGTCTCAAGAAACAAAGGAATGTGGAC 362
 QY 138 ValProLeuValIleCysGlyAsnIlyGlyAspArgPheTyrArgGluValAspGln 157
 Db 363 GTGCGCTGTGTCATTTGGCGTAAAGGGGACCGGACCTTTCACGGGAGTAGACAG 422
 QY 158 ArgGlnIleGlnIleuValGlyAspAspProGlnArgCysAlaTyrPheGluIleSer 177
 Db 423 CGGAGATTGACAGCTGTGTGGTGACGACCTTACGCTTGTGCTTTCAGATCTCA 482
 QY 178 AlaIlyIysAsnSerSerLeuAspGlnMetPheArgAlaLeuPheAlaMetAlaIlyLeu 197
 Db 483 GCGAAGAAAGACAGAGCTTGACCAAGATGTTCTGTGGCTCTTTCATGCGCAACTG 542
 QY 198 ProSerGluMetSerProAspLeuHisArgIys--ValSerValGlnTyrCys--AspValI 217
 Db 543 CTTACGAGATGAGGCCCGGACCTTGCACCGCAAGGGGTATCTGTGAGATCTGGCAGCGTAC 602
 QY 217 euHisIlyValAlaLeuArgAsnIlyIlyLeuLeuArgAlaGlyIleSer 232
 Db 603 TGCACAAAGAGCTTGAGAAACAAAGACTTTCGCTGCGGGCAGC 649

RESULT 9
 LOCUS BM543630 1023 bp mRNA linear EST 20-FEB-2002
 DEFINITION AGENCOURT_6492527 NIH_MGC_124 Homo sapiens cDNA clone IMAGE:5726764
 5', mRNA sequence.
 ACCESSION BM543630
 VERSION BM543630.1 GI:18774186
 KEYWORDS EST.


```

SOURCE          human.
ORGANISM        Homo sapiens
REFERENCE       Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS        Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE          1 (bases 1 to 1023)
JOURNAL        NIH-MGC http://mgs.nci.nih.gov/.
COMMENT        National Institutes of Health, Mammalian Gene Collection (MGC)
               Unpublished (1999)
               Contact: Robert Strausberg, Ph.D.
               Email: cgapbs-re@mail.nih.gov
               Tissue Procurement: Invitrogen
               cDNA Library Preparation: Life Technologies, Inc.
               cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
               DNA Sequencing by: Agencourt Bioscience Corporation
               Clone distribution: MGC clone distribution information can be
               found through the I.M.A.G.E. Consortium/LLNL at:
               http://image.llnl.gov
               Plate: LLAM12793 row: m column: 05
               High quality sequence stop: 637.

FEATURES       Location/Qualifiers
               1..1023
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /clone="IMAGE:5726764"
               /clone_lib="NIH_MGC_124"
               /tissue_type="hippocampus"
               /lab_host="DH10B"
               /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: EcoRV
               (destroyed); Site_2: NotI; RNA source male hippocampus,
               age 27. Library is oligo-dT primed and directionally
               cloned (EcoRV site is destroyed upon cloning). Average
               insert size 1.4 kb, insert size range 0.9-4 kb. Library is
               normalized and enriched for full-length clones and was
               constructed by C. Gruber (Invitrogen). Research Genetics
               tracking code 012."
BASE COUNT     196 a 378 c 281 g 168 t
ORIGIN
Alignment Scores:
Pred. No.:      7,13e-120      Length:      1023
Score:          1038.00        Matches:    222
Percent Similarity: 83.46%      Conservative: 5
Best Local Similarity: 81.62%    Mismatches: 30
Query Match:    71.59%         Indels:     15
DB:             13             Gaps:       3

US-09-709-103-1F1 (1-282) x BM543630 (1-1023)

QY 1 MetLysLeuAlaAlaMetIleLysLysMetCysProSerAspSerGluLeuSerIlePro 20
Db 149 ATGAACCTGGCCGGGATGATCAAGAAGATGTCCTCGGCTCGTCCGAGCTCGGAGCTGAGTATCCCG 208
QY 21 AlalysAsnCysTyrArgMetValIleLeuGlySerSerLysValGlyLysThrAlaIle 40
Db 209 GCCAAGAACTGCTATCGATGATCATCTCGGCTCGTCCGAGGTGGCAGAGCGGCATC 268
QY 41 ValSerArgPheLeuThrGlyArgPheGluAspAlaTyrThrProThrIleGluAspPhe 60
Db 269 GTGTCGGCTTCCTCAGCGGGCGGCTTCAGGAGCGCTTACACGCTACCATCGAGGACTTC 328
QY 61 HisArgLysPheTyrSerIleArgGlyGluValTyrGlnLeuAspIleLeuAspThrSer 80
Db 329 CACCGCAAGTTTCTACTCATCCGGGGGAGGTCTTACCAGCTCGACATCTCTCGACACGTCC 388
QY 81 GlyAsnHisProPheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIle 100
Db 389 GGCAACCAACCGTTCCTCCGCGCATGCGCGGCTCTCCATCTCTCACAGGAGCGTTTTCATC 448
QY 101 LeuValPheSerLeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGlnGln 120
Db 449 CTGGTGTTTCAGTCTGGACAAACCGCGACTCTCTCAGGAGGTGAGCGGCTCAGGCAGCAG 508
QY 121 IleLeuAspThrLysSerCysLeuLysAsnLysThrLysGluValAsnValAspValProLeu 140

```

```

Db 509 ATCTCTCGACACCAAGTCTTGCTCTCAAGAACAAACCAAGAGAACTGGACGTGCCCCCTG 568
QY 141 ValIleCysGlyAsnLysGlyAspArgAspPheTyrArgGluValAspGlnArgGluIle 160
Db 569 GTCACTCTGGCGCAACAAGGTGACCGGACTTCTACCGGAGGTGACCGAGGATC 628
QY 161 GluGlnLeuValGlyAspAspProGlnArgCysAlaTyrPheGluIleSerAlaLysLys 180
Db 629 GAGCAGCTGTGTGGCGACGACCCCGGCGCTCTTCTCGAGATCTCGGCCAAGAAG 688
QY 181 AsnSerSerLeu-AspGlnMetPheArgAlaLeuPheAlaMetAla-LysLeu-ProSer 199
Db 689 AACAGCAGCTCTGGGACCAAGATGTTCCGGCGCTCTTCCCATGGCCCAAGTGCCTCCCGC 748
QY 200 GluMetSerProAspLeuHisArgLys---ValSerValGlnTyrCys---AspValLeu 217
Db 749 GAGATGAGCCAGACCTGCACCGCCCAAGGTCTCGGGTGCAGGTACTGCGGACGCGGTGG 808
QY 218 HisLysLysAlaLeuArgAsnLysLysLeuLeuArgAlaGlySerGlyGlyGlyGly 237
Db 809 CACAAGAAAGGCGCTGGCGGAAACACAGAAACTGGCTGTGGGGCGCGGCAATCCGCGGGG 868
QY 238 -----AspProGlyAspAlaPheGlyIleValala 247
Db 869 CCTCTCTCTAGTCATAAACCTTGACCGAAGGACCTTGGGCGACCTAGGGCCACCATTTT 928
QY 248 PropheAlaArgArgProSerValHisSer 257
Db 929 CCGGACAGTGGCGGCTCTTGTGTCATCC 958

RESULT 10
BQ067637
LOCUS      BQ067637.1 GI:19896683
DEFINITION 5', mRNA sequence.
ACCESSION  BQ067637
VERSION    BQ067637.1
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 1032)
            NIH-MGC http://mgs.nci.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-re@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12793 row: n column: 15
            High quality sequence stop: 602.

FEATURES       Location/Qualifiers
               1..1032
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /clone="IMAGE:5755214"
               /clone_lib="NIH_MGC_115"
               /lab_host="DH10B"
               /note="Organ: pooled brain, lung, testis; Vector:
               pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
               source anonymous pool of 6 male brains, age range 23-27; 1
               male lung, age 27; and 1 male testis, age 69. Library is
               oligo-dT primed and directionally cloned (EcoRV site is
               destroyed upon cloning). Average insert size 1.8 kb,
               insert size range 1-3 kb. Library is normalized and
               enriched for full-length clones and was constructed by C.

```

Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH-MGC library."
BASE COUNT 210 a 375 c 298 g 148 t 1 others
ORIGIN

Alignment Scores:
Pred. No.: 3,09e-119 Length: 1032
Score: 1033.00 Matches: 225
Percent Similarity: 88.37% Conservative: 3
Best Local Similarity: 87.21% Mismatches: 15
Query Match: 71.24% Indels: 15
DB: 14 Gaps: 3

US-09-709-103-1f1 (1-282) x BQ067637 (1-1032)

```

QY 1 MetLysLeuAlaAlaMetLLeLysLysMetCysProSerAspSerGluLeuSerllePro 20
Db 201 ATGAAATCGGCGCGCATATGATCAAGAAATGTCGCCGACGACTCGAGCTAGATATCCCG 260
QY 21 AAlaLysAenCysTyrArgMetVallleuLgYserSerLysValGlyLysThrAlaIle 40
Db 261 GCCAAGAACTGATCGATGTCATCTCCGCTGTCCTCAAGTGGCAAGACGCGCATC 320
QY 41 ValSerArgPheLeuThrGlyArgPheGluAspAlaTyrThrProThrIleGluAspPhe 60
Db 321 GTGTGCGCTTCTTACCGCGCGCTTCAAGAGACGCTTACACGCTTACATCGAGACTTC 380
QY 61 HisArgLysPheTyrSerlleArgGlyLysValTyrGlnLeuAspIleLeuAspThrSer 80
Db 381 CACCGCAAGTTCTACTCATCCGCGCGAGGCTTACAGCTTCGACATCTCGACACGCTTC 440
QY 81 GlyAsnHisProPheProAlaMetArgArgLeuSerlleLeuThrGlyAspValPheIle 100
Db 441 GGCACACCACTCCGCTCCCGCATGCGCGCTTCTTCACTTCAAGAGAGAGCTTTTCATC 500
QY 101 LeuValPheSerLeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGln 120
Db 501 CTGGGTTCAGCTGAGCAACCGCGACTCTTCAGAGAGGTGACGCGCTTCAAGCAGCAG 560
QY 121 IleLeuAspThrLysSerCysLeuLysAsnLysThrLysGluAsnValAspValProleu 140
Db 561 ATCTCTGACACCAAGCTTCTGCTCAAGAACCAAGGAGGAGGAGGAGGAGGAGGAGG 620
QY 141 ValIleCysGlyAsnLysGlyAspArgAspPheTyrArgGluValAspGlnArgGluIle 160
Db 621 GTCACTCTGCGGCAACAGGAGTGCAGCTTCTTCAAGGAGGAGGAGGAGGAGGAGGAG 680
QY 161 GluGlnLeuValGlyAspAspProGlnArgCysAlaTyrPheGluIleSerAlaLysLys 180
Db 681 GAGCAGCTGTGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 740
QY 181 AsnSerSerLeuAspGlnMetPheArgAlaLeuPheAlaMetAlaLysLeuProSer- 199
Db 741 AACAGCAACTCGGACCAAGATGTTCCGCGCTTCTTCCCAAGGAGGAGGAGGAGGAGG 800
QY 200 GluMetSer-ArgAspLeuHisArg-LysValSerValGln---TyrCysAspValLeuH 218
Db 801 GAGATGAGCCCAAGACCTGAGACCGGCAAGGTCTGCTGCAATTACATGCAACGCGGCTGC 860
QY 218 IeLysLys---AlaLeuArgAsnLysLysLeuLeu-----A 229
Db 861 ACAAAGAGGCGCGCTGCGGAGCAAAACCTTCTGCGGCGGCGGAGCAACCGCGGCGC 920
QY 229 rgaLagLysArgGlyGlyGlyAspProGlyAspAlaPhe 243
Db 921 GGGCGGCGCGGCAACACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 964

```

RESULT 11
BF613135
LOCUS
DEFINITION
IMAGE:3473357 5 similar to IR:035626 035626 RAS.
DEXAMETHASONE-INDUCED 1, mRNA sequence.

ACCESSION BF613135
VERSION BF613135.1 GI:11784463
KEYWORDS EST.
SOURCE African clawed frog.
ORGANISM Xenopus laevis

REFERENCE
AUTHORS 1 (bases 1 to 726)
Clifton, S., Johnson, S.L., Blumberg, B., Song, J., Hillier, L., Pape, D., Martin, J., Wylie, T., Underwood, K., Theising, B., Bowers, Y., Person, B., Gibbons, M., Harvey, N., Ritter, E., Jackson, Y., McCann, R., Waterston, R. and Wilson, R.

TITLE
JOURNAL
COMMENT
Washu Xenopus EST project, 1999
Unpublished (1999)
Other ESTs: de30603.xl
Contact: Sandy Clifton, Ph.D.
Washu Xenopus EST project, 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810

Email: est@watson.wustl.edu
Library constructed by A.M. Zorn (Wellcome/CRC Institute). DNA
Sequencing by: Washington University Genome Sequencing Center
Clone distribution: Xenopus clones from this library are available
through the I.M.A.G.E. Consortium/BLNU at: info@image.llnl.gov
Seq primer: -40RP from Gibco
High quality sequence stop: 519.

FEATURES
source
Location/Qualifiers

1..726
/organism="Xenopus laevis"
/db_xref="taxon:8355"
/clone="IMAGE:3473357"
/clone_lib="Wellcome CRC prn3 dorsal 1lp"
/csize_type="dorsal 1lp"
/lab_host="DH10B (phage-resistant)"
/note="Vector: pBSRN3; Site 1: NotI; Site 2: EcoRI; cDNAs
were oligo-dT primed and directionally cloned. Staging
according to Nienukoop and Faber. Library was constructed
by A.M. Zorn (Wellcome/CRC Institute)."

BASE COUNT 197 a 196 c 181 g 152 t
ORIGIN
Alignment Scores:
Pred. No.: 3,07e-114 Length: 726
Score: 991.50 Matches: 187
Percent Similarity: 92.89% Conservative: 22
Best Local Similarity: 83.11% Mismatches: 15
Query Match: 68.38% Indels: 1
DB: 12 Gaps: 1

US-09-709-103-1f1 (1-282) x BF613135 (1-726)

```

QY 3 LeuAlaAlaMetLLeLysLysMetCysProSerAspSerGluLeuSerlleProAlaLys 22
Db 55 CTACCGCAATGATCAAGAAATGTCGCCGACGAGGAGGAGGAGGAGGAGGAGGAGG 114
QY 23 AenCysTyrArgMetVallleuLgYserSerLysValGlyLysThrAlaIleValSer 42
Db 115 AACTGTTCACGATGTCATCTCCGCTTCTTCAAGTGGCAAGACCTCCATCTGTGCT 174
QY 43 ArgPheLeuThrGlyArgPheGluAspAlaTyrThrProThrIleGluAspPheHisArg 62
Db 175 CGCTTCTTAACGCGCGATTCGAGCAATACGACGACGACGACGACGACGACGACGACG 234
QY 63 LysPheTyrSerlleArgGlyLysValTyrGlnLeuAspIleLeuAspThrSerGlyAsn 82
Db 235 AAGTTTACAGCATTCGGGAGAGGCTTACAGCTGAGATATCTTGACACCTCGGCGAAC 294
QY 83 HisProPheProAlaMetArgArgLeuSerlleLeuThrGlyAspValPheIleLeuVal 102
Db 295 CACCCCTTCCCGCATGAGAACTTCCATCTTACCGGAGATGTCATCTTGGTC 354

```

```

Qy 103 PheSerLeuAspSerPheGluValGlnArgLeuArgGlnGlnLeu 122
Db 355 TTCACCTGGACACAGGACTCTTTGAGGAGGTGAAAGGCTGAAGCAGAGATTATG 414
Qy 123 AspThrLysSerCysLeuLysAsnLysThrLysGluAsnValAspValProLeuValIle 142
Db 415 GAGACCAAGTCTCTCAAGAACAAACCAAGAGACGTTGGATGTTCCCATCGTTATC 474
Qy 143 CysGlyAsnLysGlyAspArgAspPheTyrArgGluValAspGlnArgGluLeuGluGln 162
Db 475 TCGGGGAATTAAGTAGACAGGACTTCTACAGGGAAGTCAAGCTCATGAGATTGAGCAG 534
Qy 163 LeuValGlyAspAspProGlnArgCysAlaTyrPheGluIleSerAlaLysLysAsnSer 182
Db 535 CTGGTTGGAGAGGAC---AGTAAATGCTTTACTTTGAGGTGCGCCCAAGAAAGAACTC 591
Qy 183 SerLeuAspGlnMetPheAlaLeuPheAlaLeuPheAlaMetAlaLysLeuProSerGluMetSer 202
Db 592 AGCTGGATGAGATGTTTAAAGCTCTCTTACCATTGGACAAAGTTGCCAGCGAGATGAGC 651
Qy 203 ProAspLeuHisArgLysValSerValGlnTyrCysAspValLeuHisLysLysAlaLeu 222
Db 652 TCAGACCTGCACCGCAAGGTGCTGTCCAGTACTGTGAGATCTTACACAAAGAGTCCCTG 711
Qy 223 ArgAsnLysLysLeu 227
Db 712 AAAGAAAGAAAGTA 726

RESULT 12
BG711792 655 bp mRNA linear EST 08-MAY-2001
LOCUS
DEFINITION
pglin.pk009.i15 Normalized Liver Library Gallus gallus cDNA clone
pglin.pk009.i15 5' similar to gb|AA43090.1|AF239157_1 (AF239157)
DEXRAS1 (Rattus norvegicus)G, mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Gallus gallus
chicken.
EST.
BG711792.1 GI:14005742

REFERENCE
AUTHORS
Burnside,J., Morgan,R.W. and Cogburn,L.A.
TITLE
Chicken ESTs from a normalized liver library
JOURNAL
Unpublished (2001)
COMMENT
Contact: Joan Burnside
Molecular Endocrinology
University of Delaware
40 Townsend Hall, Newark, DE 19717, USA
Tel: 302 831-1345
Fax: 302-831-3411
Email: joan@udel.edu, www.chickest.udel.edu.

FEATURES
source
1..655
/organism="Gallus gallus"
/db_xref="taxon:9031"
/clone="pglin.pk009.i15"
/clone_lib="Normalized Liver Library"
/sex="Male and Female"
/tissue_type="liver"
/lab_host="E.coli EMDH10B"
/note="vector: pCMVSPORT 6"

BASE COUNT 145 a 206 c 184 g 104 t 16 others
ORIGIN

Alignment Scores:
Pred. No.: 9.68e-114 Length: 655
Score: 987.00 Matches: 192
Percent Similarity: 97.12% Conservative: 10
Best Local Similarity: 92.31% Mismatches: 5
Query Match: 68.07% Indels: 1

```

```

DB: 12 Gaps: 0
US-09-709-103-1f1 (1-282) x BG711792 (1-655)

Qy 1 MetLysLeuAlaMetIleLysLysMetCysProSerAspSerGluLeuSerIlePro 20
Db 14 ATGAACCTGGCAGCGATGATCAAGAAGATGTGTCCAGCAGGCTGAGCTGAGCATCCCC 73
Qy 21 AlalysAsnCysTyrArgMetValIleuGlySerSerLysValGlyLysThrAlaIle 40
Db 74 GCCAAGAACTCTCCCGCATGGTTCATCTGGGCTCTCCCAAGGTGGGCAAGACGGCCATC 133
Qy 41 ValSerArgPheLeuThrGlyValArgPheGluAspAlaTyrThrProThrIleGluAspPhe 60
Db 134 GTCTCGCGCTTCTTCACCGCGGCTTCGAGGAGCAGTACACGCCCAACATCAGAGACTTC 193
Qy 61 HisArgLysPheTyrSerIleArgGlyGluValTyrGlnLeuAspIleLeuAspThrSer 80
Db 194 CACCGCAAGTTTACAGCATCCGCGGTGAGGTCTACCAGGTTCGACATCTCTGGACACGCTG 253
Qy 81 GlyAsnHisProPheProAlaMetArgLeuSerIleLeuThrGlyAspValPheIle 100
Db 254 GGCACACCACTTCCCGCATCGCGGCTGTCCATCTCCACAGGTGACGTTTTCATC 313
Qy 101 LeuValPheSerLeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGlnGln 120
Db 314 CTCGTGTTTCAGCTTGGACACCGGAGCTCTCTTCGAGGAGGTGCGAGCGCTGAAGCAGCAG 373
Qy 121 IleLeuAspThrLysSerCysLeuLysAsnLysThrLysGluAsnValAspValProLeu 140
Db 374 ATCTCTGGAGACCAAGTCTGCTGAAGAACAAACCAAGGAGAACATCGAGGTCCGCTG 433
Qy 141 ValIleCysGlyAsnLysGlyAspArgAspPheTyrArgGluValAspGlnArgGluIle 160
Db 434 GTCATCTCGGCAACAAAGCGGCGGAGCTTTACCGGGAGGTGCGAGCCCGCAGAGATC 493
Qy 161 GluGlnLeuValGlyAspAspProGlnArgCysAla-TyrPheGluIleSerAlaLysLY 180
Db 494 GAGCAGCTGTGGCGCAGACCCCAAGAAATCGCANNACTTCGAGATCTCGGCCNAGAA 553
Qy 180 sAsnSerSerLeuAspGlnMetPheArgAlaLeuPheAlaMetAlalysLeuProSerGI 200
Db 554 GAACAGCAGCTGGATCAGATGTTCCAGCGCTCTTCGCCATGCGCAAACTGCCAGCGGA 613
Qy 200 uMetSerProAspLeuHisArg 207
Db 614 GATGAGCCCGACCTGCACCGC 635

RESULT 13
BG711792 655 bp mRNA linear EST 05-SEP-2001
LOCUS
DEFINITION
603190722P1 NIH_MGC_95 Homo sapiens cDNA clone IMAGE:5262101 5',
mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1 (bases 1 to 758)
NIH-MGC http://mgi.nci.nih.gov/.
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
Contact: cgabbs-r@mail.nih.gov
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:

AUTHORS
TITLE
JOURNAL
COMMENT
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)

```

http://image.llnl.gov
Plate: LLM11660 row: 1 column: 06
High quality sequence stop: 738.
Location/Qualifiers

FEATURES

source

1..758
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5262101"
/clone_1ib="NIH MGC 95"
/tissue_type="hippocampus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescript (modified pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcag 5' Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3', size-selected for average insert size 2.5 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIMH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."

BASE COUNT 155 a 252 c 249 g 102 t
ORIGIN

Alignment Scores:

Pred. No.: 6e-106 Length: 758
Score: 926.00 Matches: 182
Percent Similarity: 98.92% Conservative: 1
Best Local Similarity: 98.38% Mismatches: 1
Query Match: 63.86% Indels: 1
DB: 13 Gaps: 0

US-09-709-103-1f1 (1-282) x B1553776 (1-758)

QY 98 ValPheileLeValPheSerLeuSpaenAArgAspSerPheGluGluValGlnArgLeu 117
DB 6 GTTTCATCTCTGTGTTCTGCTGCAACCGCGACTCTTGGAGAGGTGCGGCTC 65
QY 118 ArgGlnGlnLeuLeuAspThrLysSerCysLeuLysAsnLysThrLys-GluAsnValas 137
DB 66 AGCGACGAGATCTCGACCAACCAAGTCTGCTCAAGAACAAACCAAGGAGAACTGGA 125
QY 137 PValProLeuValLleCysGlyAsnLysGlyAspArgAspPheTyrArgGluValAspG1 157
DB 126 CGTGGCCCTGTCATCTGGGCGCAACAGGTGACCGCGACTTCTACCGCGAGGTGACCA 185
QY 157 nArgGlnLeuGlnLeuValGlyAspAspProGlnArgCysAlaTyrPheGlnLleSe 177
DB 166 GCGCGAGATCGACGAGCTGTGGGCGACGACCCCAAGGCTGCGCTTCTGAGATCTC 245
QY 177 rAlaLysLysAsnSerSerLeuAspGlnMetPheArgLaleuPheAlaMetAlaLysLe 197
DB 246 GCGCAAGAAAGACAGACAGCTGGACCAAGTGTCCGCGGCTCTTGGCATGGCCAAAGCT 305
QY 197 nProSerGlnMetSerProAspLeuHisArgLysValSerValGlnTyrCysAspValle 217
DB 306 GCGCGACGAGAGACCCAGACTGACCGCAAGGTCTCGGTGCACTGCGACGTGCT 365
QY 217 whisLysLysAlaLeuArgAsnLysLysLeuLeuArgLaleGlySerGlyGlyGlyG1 237
DB 366 GCAAGAAAGGCGCTGCGAAGAGTGTGGCGGCGGCGGCGGCGGCGGCGGCGG 425
QY 237 yAspProGlyAspAlaPheGlyLleValAlaProPheAlaArgArgProSerValHisSe 257
DB 426 CGACCCGCGCGCGCTTGGCATGTGGCACTCTCGCGCGCGCGCGCGCGCGCTACAG 485
QY 257 rAspLeuMetTyrLleArgGluLysAlaSerAlaGlySerGlnAlaLysAspLysGluAr 277
DB 486 CGACCTCAATGATCCCGGAGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 545
QY 277 gCyValLleSer 281
DB 546 CTGCTCATCAGC 558

RESULT 14
BM921656 1137 bp MRNA linear EST 12-MAR-2002
LOCUS BM921656
DEFINITION AGENCOURT_6708111 NIH_MGC_115 Homo sapiens CDNA clone IMAGE:5753182
5', mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM12788 row: 1 column: 23
High quality sequence stop: 542.
Location/Qualifiers

FEATURES

source

1..1137

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:5753182"

/clone_1ib="NIH MGC_115"

/lab_host="DH10B"

/note="Organ: pooled brain, lung, testis; Vector:
pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 Kb,
insert size range 1-3 Kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH MGC Library."

BASE COUNT 223 a 427 c 315 g 172 t
ORIGIN

Alignment Scores:

Pred. No.: 9.93e-105 Length: 1137
Score: 918.50 Matches: 212
Percent Similarity: 75.25% Conservative: 10
Best Local Similarity: 71.86% Mismatches: 45
Query Match: 63.34% Indels: 29
DB: 14 Gaps: 3

US-09-709-103-1f1 (1-282) x BM921656 (1-1137)

QY 1 MelLysLeuAlaLaleLleLysLysMetCysProSerAspSerGlnLeuSerLlePro 20
DB 221 ATGAACTGCGCCCGATGATCAAGAGATGTGCGCCGAGGAGCTCGAGCTGATCCCG 280
QY 21 AlLysAsnCysTyrArgMetValLleLeuGlySerSerLysValGlyLysThrAlaLe 40
DB 281 GCAAGAACTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 340
QY 41 ValSerArgPheLeuThrGlyArgPheGluAspAlaTyrThrProThrLleGlnAspPhe 60
DB 341 GTGTGGGCTTCTTCCACCGCGCTTCCAGGAGCGCTTACCGCTTACATCGAGGACTTC 400
QY 61 HisArgLysPheTyrSerLleArgGlyGluValTyrGlnLeuAspLleLeuAspThrSer 80
DB 401 CACCGAAGTTTACTTACATCCGCGCGGCGAGGTCTTACCACTGAGACATCTCTCGACGCTCC 460

```

Qy 81 GlyAsnHisProPheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPhe11e 100
Db 461 GGCAACACCCGTTCCCGCGCATGCGCGCTCTCCATCCTCACAGGAGACGTTTTTCATC 520

Qy 101 LeuValPheSerLeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGlnGln 120
Db 521 CTGGTGTTCAGTCTGGACAAACCGGAGCTCTCTTCAGGAGGTGCGCGCTCAGCGACGAG 580

Qy 121 IleLeuAspThrLysSerCysLeuLysAsnLysThrLysGluAsnValAspValProLeu 140
Db 581 ATCTCTCACACCAAGTCTTGCTTCAAGAACAAACCAAGGAGATGTGGAGCGTGCCTTG 640

Qy 141 ValIleCysGlyAsnLysGlyAspArgAspPheTyrArgGluVal-AspGlnArgGlu11 160
Db 641 GTCATCTCGGCAACACAGAGTGACCGGAGCTTCTACCGCGAGGTGGGACAAACCGGAGAT 700

Qy 160 eGluGlnLeuValGlyAspAspProGlnArgCysAlaTyrPheGluIleSerAlaLysLy 180
Db 701 CGAGCAGCTGGTGGGCGG-ACACCCACGCGGTGGCGCTCTCTCGAGATCTCGGCCAAGAA 759

Qy 180 sAsnSer-SerLeuAsp-GlnMetPheArgAlaLeuPheAlaMetAlaLys-LeuProSe 199
Db 760 AAACAGCATCTCTGACCCCAATGTTCCGCGGGCTCTTCCCTGGCCCAACCTGCCAG 819

Qy 199 r-----GluMetSerProAspLeuHisArgLysValSerValGlnTyrCy 214
Db 820 CGGAGAACGAGTCCCGGAACCTTGGACCCCAAGTCTCCGCGGGGAATAACTGGCGAAC 879

Qy 214 sAspValLeuHisLysLysAlaLeuArgAsnLysLysLeuLeuArg-AlaGlySerGlyG 234
Db 880 GGCTTGGCCCCATAAAGGGCGCTGCCGAAACAAAGAAACTGGCTTGGCGGACCGGCC 939

Qy 234 lYgLYgLYgLYgPro----- 239
Db 940 ACCGGCTGGTGGCGGCTCTGTCACCGCGGGGGAACGCTCTCTTGTGGCGAACCC 999

Qy 240 --GlyAsp-----AlaPheGlyLeValAlaProPheAlaArgArgProSerValH 256
Db 1000 GGGGGGACACCTTGTTCGCGCGCGGCGTGCACCAATGGTACAAACCGGCCAACCCCTC 1059

Qy 256 isSerAspLeuMetTyrIleArgGluLysAlaSer 267
Db 1060 CCCGGTTACACTCTCTCTCCGCAAAAGGGTCC 1094

RESULT 15
BI596688 814 bp mRNA linear EST 07-SEP-2001
LOCUS 603243262F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5285894 5',
DEFINITION mRNA sequence.
ACCESSION BI596688
VERSION BI596688.1 GI:15489627
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 814)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11722 row: k column: 15
High quality sequence stop: 803.
Location/Qualifiers
FEATURES

```

```

source 1. .814
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5285894"
/clone_lib="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/notes="Organ: brain; Vector: pBluescriptR (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
); Oligo-dr primed using primer 5'-TTTTTTTTTTTTTTVN-3',
size-selected for average insert size 2.3 kb and
normalized to Rot 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT 160 a 310 c 225 g 119 t
ORIGIN
Alignment Scores: 9.15e-105 Length: 814
Pred. No.: 917.00 Matches: 192
Score: 96.48% Conservative: 0
Percent Similarity: 96.48% Mismatches: 5
Best Local Similarity: 96.48% Indels: 4
Query Match: 63.24% Gaps: 0
Db: 13
US-09-709-103-1F1 (1-282) x BI596688 (1-814)

```

```

Qy 1 MetLysLeuAlaAlaMetIleLysLysMetCysProSerAspSerGluLeuSerIlePro 20
Db 220 ATGAAACTGGCGCGCATGATCAAGAGATGTGCCGAGGACTCGGAGTCGATATCCCG 279

Qy 21 AlaLysAsnCysTyrArgMetValIleLeuGlySerSerLysValGlyLysThrAla11e 40
Db 280 GCCAAGAACTGATATCGCATGTGTCATCTCGCTCGTCCAAGGTGGGCAAGCGCCATC 339

Qy 41 ValSerArgPheLeuThrGlyArgPheGluAspAlaTyrThrProThrIleGluAspPhe 60
Db 340 GTGTGCGCGCTTCTCACCGCGCGCTTCGAGGAGCGCTACACGCTTACCCTACGAGACTTC 399

Qy 61 HisArgLysPheTyrSerIleArgGlyGluValTyrGlnLeuAspIleLeuAspThrSer 80
Db 400 CACCGCAAGTTCTACTCCATCCGCGCGGAGGTCTACACGCTCGACATCTCTGACACGTC 459

Qy 81 GlyAsnHisProPheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPhe11e 100
Db 460 GGCAACACCCGTTCCCGCGCATGCGCGCTCTCCATCCTCACAGGAGACGTTTTTCATC 519

Qy 101 LeuValPheSerLeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGlnGln 120
Db 520 CTGGTGTTCAGTCTGGACAAACCGGAGCTCTTCGAGGAGGTGCGCGCTCAGGCGAG 579

Qy 121 IleLeuAspThrLysSerCysLeuLysAsnLysThrLysGluAsnValAspValProLeu 140
Db 580 ATCTCTCACACCAAGTCTTGCTTCAAGAACAAACCAAGGAGAACGTGGAGCGTGCCTTG 639

Qy 141 ValIleCysGlyAsnLysGlyAspArgAspPheTyrArgGluValAspGlnArgGlu11e 160
Db 640 GTCATCTCGGCAACAAAGGAGTTCACCGGAGCTTCTTACCGCGAGGTGGACCGCGGAGATC 699

Qy 161 GluGlnLeuValGlyAspAsp-ProGlnArgCysAlaTyrPheGlu-IleSerAlaLysL 180
Db 700 GAGCAGCTGTGGGGGAGACCCCGGAGCGCTGCGGCTTCTTCGAGATCTTCGCGCAAGA 759

Qy 180 YsAsnSerSerLeuAspGlnMetPheArgAlaLeuPheAlaMetAlaLysLeu 197
Db 760 AGACAGCAGCGCG-GACACAGATGTC-CGCGCGCTCTTTCGCCCATGGCAAGCTG 810

```

Search completed: December 30, 2002, 16:16:36
Job time : 2198 secs

SUMMARIES

2079.951 Million cell updates/sec

Scoring table:

Xgapop	10.0	Xgapext	0.5
Ygapop	10.0	Ygapext	0.5
Fgapop	6.0	Fgapext	7.0
Delop	6.0	Delext	7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Command line parameters:

```

positional_argument(model="DEV=xdp
-Q/cgn2.1/JSPID_spool/US09J09103/runat_20122002.162406.818/bpp_query_fasta_1.455
-DB-EST -Qfmt=fastest -SUFFIX=220.ist -MINNAK=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsme2 -TRANS=human40.cdi -LIST=45
-LOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=15 -MODE=LOCAL
-OUTPMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=300000000
-USER=US09J09103_@cgn_1.1716. @runat_20122002.162406.818 -NCPG=6 -ICP=3
-NO_XLXPY -NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THRRAU=1 -XGAP=10 -XGAPEXT=0.5 -FGAP=6 -FGAPEXT=7
-XGAP=10 -XGAPEXT=0.5 -DELOP=6 -DELEXT=7

```

Database : EST:★

```

1: em_estbm: *
2: em_esthm: *
3: em_estln: *
4: em_estmu: *
5: em_estov: *
6: em_estpl: *
7: em_estro: *
8: em_htc: *
9: gb_est1: *
10: gb_est2: *
11: gb_htc: *
12: gb_est3: *
13: gb_est4: *
14: gb_est5: *
15: em_estfun: *
16: em_estom: *
17: gb_gss: *
18: em_gss_hnu: *
19: em_gss_inv: *
20: em_gss_pln: *
21: em_gss_vtr: *
22: em_gss_fun: *
23: em_gss_man: *
24: em_gss_mus: *
25: em_gss_other: *
26: em_gss_pro: *
27: em_gss_rtd: *

```

ALIGNMENTS

Unit No.	Score	Query Match	Length	DB	ID	Description
1	1210.5	83.5	1035	14	BM919341	BM919341 AGENCOCURT
2	1165	80.4	962	13	BM543472	BM543472 AGENCOCURT
3	1144	79.0	1103	14	BM9200514	BM9200514 AGENCOCURT
4	1121	77.4	965	9	AL533318	AL533318
5	1120	77.3	742	10	AM028127	AM028127
6	1114.5	76.9	958	14	BM0719566	BM0719566 AGENCOCURT
7	1100.5	75.9	1053	14	BM9217377	BM9217377 AGENCOCURT
8	1064	73.4	649	12	BG0850900	BG0850900 H3108E05-
9	1038	71.6	1023	13	BM543630	BM543630 AGENCOCURT
10	1033	71.3	1032	14	BQ0676377	BQ0676377 AGENCOCURT
11	991.5	68.4	726	12	BF613135	BF613135 de30e03.y
12	987	68.1	655	12	BG711722	BG711722 p91in.pk0
13	926	63.9	758	13	BM153776	BM153776 603190722
14	918.5	63.4	1137	14	BM921656	BM921656 AGENCOCURT
15	917	63.3	814	13	B1595668	B1595668 603243262
16	896.5	61.9	1160	14	BM805574	BM805574 AGENCOCURT
17	892	61.6	644	13	BG969048	BG969048 602835003
18	890.5	61.5	831	13	B1596637	B1596637 603243203
19	885.5	61.1	710	13	B1526038	B1526038 BJS26038
20	884	61.0	1006	14	BQ073742	BQ073742 AGENCOCURT
21	869	60.0	506	13	BM111047	BM111047 1g60a06.y
22	867	59.8	1300	11	AK015898	AK015898 MS_muscu
23	862	59.5	947	14	BQ954076	BQ954076 AGENCOCURT
24	859	59.3	648	10	B8636889	B8636889
25	859	59.3	689	10	BB632659	BB632659
26	854	58.9	904	14	BQ947936	BQ947936 AGENCOCURT
27	830	58.7	617	13	B1393659	B1393659 p9p1n.pk0
28	832.5	57.5	640	13	B1490945	B1490945 BJS46945
29	825	56.9	699	12	BG706012	BG706012 602669154
30	812	56.0	699	12	BG709229	BG709229 602674515
31	812	56.0	700	13	B1595609	B1595609 603243558
32	809	55.8	589	13	BM426006	BM426006 p9t2n.pk0
33	807.5	55.7	888	13	B1754083	B1754083 603027639
34	785.5	54.2	1332	13	BM460899	BM460899 AGENCOCURT
35	767	52.9	674	13	B1601563	B1601563 603250773
36	748	51.6	675	13	B1545172	B1545172 603242791
37	745	51.4	666	13	B1549939	B1549939 603194786
38	741	51.1	571	13	B1682222	B1682222 464129 MA
39	732.5	50.6	456	12	BF555582	BF555582 UT-R-E0-C
40	730	50.4	509	10	AW915326	AW915326 EST346630
41	703.5	48.6	505	13	B1591213	B1591213 p9p1n.pk0
42	685	47.3	542	9	AA790463	AA790463 v0M4G12.r
43	674.5	46.5	1438	14	BM807669	BM807669 AGENCOCURT
44	660.5	45.6	435	10	AP990252	AP990252 ut35911.y
45	655.5	45.2	461	13	BT489219	BT489219

RESULT 1	
BM919341	
LOCUS	
DEFINITION	BM919341 1035 bp mRNA linear EST 12-MAR-2002
ACCESSION	AGNCOCURF_6715661 NIH_MGC_120 Homo sapiens cDNA IMAGE:5748585
VERSION	5', rRNA sequence.
KEYWORDS	BM919341 BM919341.1 GI:19369720
SOURCE	EST.
ORGANISM	human.
	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE	1 (bases 1 to 1035)
AUTHORS	NIH-MGC http://mgc.ncl.nih.gov/ .
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: c9abbs-remail.nih.gov

Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: L1AM12776 row: 1 column: 10
 High quality sequence stop: 658.

FEATURES

source Location/Qualifiers
 1. 1035

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_image="5748585"
 /clone_id="NIH_MGC_120"
 /lab_host="DH10B"

/note="Organ: pooled pancreas and spleen; Vector:
 PCMV-SPORET6; Site.1: NotI; Site.2: EcoRV (destroyed); RNA
 source anonymous pool of spleen and pancreas from 28 yo
 male. Library is oligo-dT primed and directionally cloned
 (EcoRV site is destroyed upon cloning). Average insert
 size 1.5 kb, insert size range 1-2.5 kb. Library is
 normalized and enriched for full-length clones and was
 constructed by C. Gruber (Invitrogen). Research Genetics
 tracking code 025. Note: this is a NIH MGC library."

BASE COUNT 209 a 386 c 293 g 146 t 1 others
 ORIGIN

Alignment Scores:

Pred. No.: 1,27e-141 Length: 1035
 Score: 1210.50 Matches: 244
 Percent Similarity: 95.028 Conservative: 4
 Best Local Similarity: 93.498 Mismatches: 8
 Query Match: 83.548 Indels: 5
 14 Gaps: 2

US-09-709-103-2 (1-281) x BM919341 (1-1035)

QY 1 MettysleuAlaMetIleLysMetCysProSerAspSerGluLeuSerIlePro 20
 Db 187 ATGAAGACTGGCCGGATGATCAAGAAAGATGTGCCGACACCTGGAGCTGATATCCG 246
 QY 21 AlalysasnCysTyrArgMetValIleLeuGlySerSerIleValGlyThrAlaIle 40
 Db 247 GCCAAGACTGCTATCCATGCTATCTCTGCTGCTCCCAAGGTGGCAAGAGCCCATC 306
 QY 41 ValSerArgPheLeuThrGlyArgPheGluAspAlaTyrThrProThrIleGluAspPhe 60
 Db 307 GTGTCGGCGCTTCTCACCGCGCGCTTCGAGAGCGCTACACGCTACCATCGAGCACTTC 366
 QY 61 HisArgLysPheTyrSerIleArgGlyValTyrGlnLeuAspIleLeuAspThrSer 80
 Db 367 CACCGCAAGTCTTCACTCATCCGCGGAGGCTTACCATCGACATCTCGACACGCTCC 426
 QY 81 GlyAsnHisProPheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIle 100
 Db 427 GGCACACCAACCGTCCCGCCATCGCGCTCTCCATCTCATACAGAGAGCGTTTTCATC 486
 QY 101 LeuValPheSerLeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGlnGln 120
 Db 487 CTGCTGTTCAGCTGGACACCGCGACTCTCTCGAGAGGTGCGAGCGCTCAGGCAACAG 546
 QY 121 IleLeuAspThrLysSerCysLeuLeuAsnLysThrLysGluAsnValAspValProLeu 140
 Db 547 ATCTCTCAACACCAAGTCTTGCCTCAAGAAACAAAGGAGACGTGGAGCGCCCTG 606
 QY 141 ValIleCysGlyAsnLysGlyAspArgAspPheTyrArgGluValAspGlnArgGluIle 160
 Db 607 GTATCTGCGGCAACAAAGGTGAGACCGGCACTTCTACGCGAGGTGGAGCAAGCGGAGATC 666

QY 161 GluGlnLeuValGlyAspAspProGlnArgCysAlaTyrPheGluIleSerAlaLysLys 180
 Db 667 GAGACAGGTGTGGGCGAGACAGCCCGAGCGCTGCTACTCTGAGATCTCGGCCAAGAG 726
 QY 181 AsnSerSerLeuAspGlnMetPheArgAlaLeuPheAlaMetAlaLysLeuProSerGlu 200
 Db 727 AACAGCACCCTGGAGCAAGATGTCTCGCGCTCTCCATCGGCAAGCTGCCACAGAG 786
 QY 201 MetSerProAspPheHisArgLysValSerValGlnTyrCysAspValLeuHisLysLys 220
 Db 787 ATGAGCCCAAGACTCTGACACCGCAAGCTCTCGCTGCACTGCGACGTCTGCAACAGAA 846
 QY 221 AlaLeuArgAsnLysLysLeuLeuArgAlaGlySerGly-GlyGly-AspProG 240
 Db 847 GCGCTCGGCAACAAAGTCTGCGGCGCGCAAGCGCGCGCGCGCGCGCGCGCGCGCGCG 906
 QY 240 LysAlaPhePheGlyIleValAla-----ProPheAlaArg---ArgProSerValHis 256
 Db 907 GCGCAGCGCCCTTTGGCATTGCTGGCAACCTTTTGGCGCGCGCGCGCGCGCGCGCGCG 965

RESULT 2 962 bp mRNA linear EST 20-FEB-2002
 BM543472
 LOCUS BM543472
 DEFINITION AGENCCOURT_6492614 NIH_MGC_124 Homo sapiens cDNA clone IMAGE:5726554
 5', mRNA sequence.
 ACCESSION BM543472
 VERSION BM543472.1 GI:18773895
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1 (bases 1 to 962)
 AUTHORS NIH-MGC http://imgc.ncl.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: c9abbs-remail.nih.gov
 Tissue Procurement: Invitrogen
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: L1AM12719 row: d column: 11
 High quality sequence start: 18
 High quality sequence stop: 692.

FEATURES

source Location/Qualifiers

1. 962
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_image="5726554"
 /clone_id="NIH_MGC_124"
 /tissue_type="hippocampus"
 /lab_host="DH10B"

/note="Organ: brain; Vector: PCMV-SPORET6; Site.1: EcoRV
 (destroyed); Site.2: NotI; RNA source male hippocampus,
 age 27. Library is oligo-dT primed and directionally
 cloned (EcoRV site is destroyed upon cloning). Average
 insert size 1.4 kb, insert size range 0.9-4 kb. Library is
 normalized and enriched for full-length clones and was
 constructed by C. Gruber (Invitrogen). Research Genetics
 tracking code 012."

BASE COUNT 185 a 363 c 272 g 141 t 1 others
 ORIGIN

Alignment Scores:

Pred. No.: 6.21e-136 Length: 962
 Score: 1165.00 Matches: 234
 Percent Similarity: 92.55% Conservative: 2
 Best Local Similarity: 91.76% Mismatches: 9
 Query Match: 80.40% Indels: 10


```

DB: 13 Gaps: 1
US-09-709-103-2 (1-281) x BM543472 (1-962)
QY 1 MetLysLeuAlaMetIleLysMetCysProSerAspSerGluLeuSerIlePro 20
DB 224 ATGAACCTGGCCCCGAGTCAAGAAAGATGTGCCACGACACGAGCTGACTATCCCG 283
QY 21 AlAluYsAncysTyrrArgMetValIleLeuGlySerSerLysValGlyThrAlaIle 40
DB 284 GCCAAGACAGCTGATGCAATGTCATCTCGCTGCTCCAGAGTGGGCAAGAGGCCATC 343
QY 41 ValSerArgPheLeuThrGlyArgPheGluAspAlaTyrThrProThrIleGluAspPhe 60
DB 344 GTGTCGGCTTCTCCACCGCGCTTCGAGAGAGCGCTACACGCTACACATCGAGACTTC 403
QY 61 HisArgLysPheTyrSerIleArgGlyGluValTyrGluLeuAspIleLeuAspThrSer 80
DB 404 CACCGCAAGTTCTACATCCATCCGCGGAGGTTCACCATCTCGACATCTCGACAGCTCC 463
QY 81 GlyAsnHisProPheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIle 100
DB 464 GGCACACACCCGTTCCCGGCAATCGCGCGCTCTCCATCTCCACAGAGACGTTTCATC 523
QY 101 LeuValPheSerLeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGlnGln 120
DB 524 CTGGTGTTCAGTCTGACACACCGGACCTCTCGAGAGAGGTGCACGCGCTCAGGCAAG 583
QY 121 IleLeuAspThrLysSerCysLeuLysAsnLysThrLysGluAsnValAspValProLeu 140
DB 584 ATCTCCACACCAAGTCTGCTCAAGACAAACCAAGAGAGACGTGACAGTGGCCCTG 643
QY 141 ValIleCysGlyAsnLysGlyAspArgAspPheTyrArgGlyValAspGlnArgGluIle 160
DB 644 GTCATCTGCGGCAACAAAGGAGGACCGGACTTCTACCGCGAGGTGACCAAGCGAGATC 703
QY 161 GluGlnLeuValGlyAspAspProGlnArgCysAlaTyrPheGluIleSerAlaLysLys 180
DB 704 GAGCAGCTGTGGGCGACGACGCCACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 763
QY 181 AsnSerSerLeuAspGlnMetPheArgAlaLeuPheAlaMetAlaLysLeuProSerGlu 200
DB 764 AACGACAGCTGACACAGATGTTCCGCGCTCTCCGCAATGCGCCACAGCTGCCAGGAG 823
QY 201 MetSerProAspLeuHisArgLysValSerValGlnTyrCysAspValLeuHisLysLys 220
DB 824 ATGAGCCACAGCTGACACGAGTCTCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 883
QY 221 AlaLeuArg-AsnLysLysLeuLeuArg-AlaGlySerGlyGlyGlyLysAspProG 240
DB 884 GCGCTGCGGGAACAACTGCTGCGGCGGCGGCAACCGCGGCGGCGGCGGCGCA 943
QY 240 LysAspAlaPheGlyIleValAlaProPheAlaArgArgPro 253
DB 944 ACCCG-----GGCGACAGCGCT 960

RESULT 3
BM920514 1103 bp mRNA linear EST 12-MAR-2002
LOCUS BM920514
DEFINITION AGENCOURT 6709473 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5750507
5', mRNA sequence.
ACCESSION BM920514
VERSION BM920514.1 GI:19370893
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1103)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.

```

Email: cgaabs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.lnl.gov
 Plate: LLM12781 row: 3 column: 12
 High quality sequence stop: 676.
 Location/Qualifiers

FEATURES

source

1..1103

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:5750507"

/clone_11b="NIH_MGC_122"

/lab_host="DH10B"

/note="Organ: pooled lung and spleen; Vector: pCMV-Sport6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of 24 week female spleens. Library is oligo-dT primed, and 20-22 week male spleens. Library is oligo-dT primed, and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 026. Note: this is a NIH-MGC Library."

```

BASE COUNT 230 a 402 c 314 g 156 t 1 others
ORIGIN
Alignment Scores: 3,42e-133 Length: 1103
Pred. No.: 1144.00 Matches: 246
Score: 88.69% Conserved: 5
Percent Similarity: 86.93% Mismatches: 24
Best Local Similarity: 78.95% Indels: 9
Query Match: 14 Gaps: 3
DB: 14

US-09-709-103-2 (1-281) x BM920514 (1-1103)
QY 1 MetLysLeuAlaMetIleLysMetCysProSerAspSerGluLeuSerIlePro 20
DB 201 ATGAACCTGGCCCCGAGTCAAGAAAGATGTGCCACGACACGAGCTGACTATCCCG 260
QY 21 AlAluYsAncysTyrrArgMetValIleLeuGlySerSerLysValGlyThrAlaIle 40
DB 261 GCCAAGACAGCTGATGCAATGTCATCTCGCTGCTCCAGAGTGGGCAAGAGGCCATC 320
QY 41 ValSerArgPheLeuThrGlyArgPheGluAspAlaTyrThrProThrIleGluAspPhe 60
DB 321 GTGTCGGCTTCTCCACCGCGCTTCGAGAGAGCGCTACACGCTACATCGAGACTTC 380
QY 61 HisArgLysPheTyrSerIleArgGlyGluValTyrGluLeuAspIleLeuAspThrSer 80
DB 381 CACCGCAAGTTCTACATCCATCCGCGGAGGTTCACCATCTCGACATCTCGACAGCTCC 440
QY 81 GlyAsnHisProPheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIle 100
DB 441 GGCACACACCCGTTCCCGGCAATCGCGCGCTCTCCATCTCCACAGAGACGTTTCATC 500
QY 101 LeuValPheSerLeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGlnGln 120
DB 501 CTGGTGTTCAGTCTGACACACCGGACCTCTCGAGAGAGTGCACGCGCTCAGGCAAG 560
QY 121 IleLeuAspThrLysSerCysLeuLysAsnLysThrLysGluAsnValAspValProLeu 140
DB 561 ATCTCCACACCAAGTCTGCTCAAGACAAACCAAGAGAGACGTGACAGTGGCCCTG 620
QY 141 ValIleCysGlyAsnLysGlyAspArgAspPheTyrArgGlyValAspGlnArgGluIle 160
DB 621 GTCATCTGCGGCAACAAAGGAGGACCGGACTTCTACCGCAGGTGACACGCGGAGATC 680
QY 161 GluGlnLeuValGlyAspAspProGlnArgCysAlaTyrPheGluIleSerAlaLysLys 180

```


JOURNAL
COMMENT

Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaabs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmett-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/biopr/image/image.html
Seq primer: -40UP from Gibco
High quality sequence stop: 445.
Location/Qualifiers

FEATURES
source

1..742
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2530668"
/clone_1ib="NCI_CGAP_Kid11"
/lab_host="DH10B"
/note="Organ: kidney; Vector: pT7AD-Pac (Pharmacia) with
a modified polylinker. Site.1: Not I; Site.2: Eco RI;
Plasmid DNA from the normalized library NCI_CGAP_Kid3 was
prepared, and ss circles were made in vitro. Following HAP
purification, this DNA was used as tracer in a subtractive
hybridization reaction. The driver was PCR-amplified cDNAs
from a pool of 5,000 clones made from the same library
(cloneids 1322376-1323911, 1456007-1456775, and
1500552-1502855). Subtraction by Bento Soares and M.
Fatima Bonaldo."

BASE COUNT 149 a 260 c 208 g 122 t 3 others
ORIGIN

Alignment Scores: 1.97e-130 Length: 742
Pred. No.: 1120.00 Matches: 220
Score: 95.748 Conservative: 5
Percent Similarity: 93.628 Mismatches: 10
Best Local Similarity: 77.29% Indels: 0
Query Match: 10 Gaps: 0

US-09-709-103-2 (1-281) x AM028127 (1-742)

QY 1 MetLysLeuAlaMetLysLysMetCysProSerAspSerGluLeuSerIlePro 20
DB 35 ATGAAGCTGGCCGCGATGATCAAGAGATGTGCCCCGAGCGACTGCGATATCCCG 94
QY 21 AlAlaYsAnCysTyrArGMeTValIleLeuGlySerSerLysValGlyLysThrAla 40
DB 95 GCCAAGAACTGCTATGCCATGTCATCGCTCGTCCCAAGTGGGCAAGAGCGCCATTC 154
QY 41 ValSerArgPheLeuThrGlyArgPheGluAspAlaTyrThrProThrIleGluAspPhe 60
DB 155 GTGTGGCTTCCTCCACCTGCGGCTTGAGAGACGCTTACACGCTTACATCGAGGACTTC 214
QY 61 HisArgLysPheTyrSerIleArgGlyGluValTyrGlnLeuAspIleLeuAspThrSer 80
DB 215 CACCCCAAGTCTACTCCATCCGCGGCGAGTACACAGCTCCGACATCCCGACAGCTTC 274
QY 81 GlyAsnHisProPheProAlaMetArGArgLeuSerIleLeuThrGlyAspValPheIle 100
DB 275 GGCACACACCCGTCCTCCGCGCATGCGGCGCTTCCATCTCCACAGAGACGTTTTCATC 334
QY 101 LeuValPheSerLeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGln 120
DB 335 CTGGTTTCAAGTCTGACACACCGGAGCTCTTCGAGGAGGTGCGCGCTCAGGACAG 394
QY 121 IleLeuAspThrLysSerCysLeuLysAsnLysThrLysGluAsnValAspValProLeu 140
DB 395 ATCTCGACACCAAGTCTTGGCTCAAGAACCAAGGAGAAAGAGTGGACGTGCCCTG 454
QY 141 ValIleCysGlyAsnLysGlyAspArgAspPheTyrArgGluValAspGlnArgGluIle 160

DB 455 GTCACTGCGGCAACAAAGGTGACCGGACTTCTACCGGAGTGGACCGGAGATC 514
QY 161 GluGlnLeuValGlyAspAspProGlnArgCysAlaTyrPheGluIleSerAlaLysLys 180
DB 515 GAGCAGCTGGGGGCGAGCAGACCCCGAGCGGTGGCTTACTTGAGATCTCNCAGAGAG 574
QY 161 AsnSerSerLeuAspGlnMetPheArgAlaLeuPheAlaMetAlaLysLeuProSerIle 200
DB 575 AACACAGACCTTGACACAGATGTTCCGGCGCTCTGCGCATGGCCAAAGCTCCAGCAG 634
QY 201 MetSerProAspLeuHisArgLysValSerValGlnTyrCysAspValLeuHisLysLys 220
DB 635 ATGAGCCCAAGACTTCACCGCAGAGCTCTCNGTGCAGTACCTGACCTGCTGCACAGAG 694
QY 221 AlaLeuArgAsnLysLysLeuLeuArgAlaGlySerGlyGly 235
DB 695 GCGCTGCGAAGACAGAACGCTGTGCGGTCCGACAGCAGCTGGT 739

RESULT 6
LOCUS B0719566 958 bp mRNA linear EST 16-JUL-2002
DEFINITION AGENCOURT.8305191 lupsk1_sympathetic_trunk Homo sapiens CDNA clone
IMAGE:6193170 5', mRNA sequence.
ACCESSION. B0719566
VERSION. B0719566.1 GI:21858463
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 958)
AUTHORS NIH-MGC http://mgs.nci.nih.gov/.
TITLES National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
CONTACT: Robert Strausberg, Ph.D.
Email: cgaabs-r@mail.nih.gov
Tissue Procurement: Dr. James R. Lupski
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13596 row: f column: 19
High quality sequence stop: 514.
Location/Qualifiers

FEATURES
source

1..958
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6193170"
/clone_1ib="lupsk1_sympathetic_trunk"
/sex="male"
/tissue.type="sympathetic trunk"
/dev_stage="adult, 16 yr"
/lab_host="DH10B"
/note="Vector: pCMV-Sport6 (Life Technologies); Site.1:
NotI; Site.2: SalI; CDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TCGACCAAGCGGCGG-3' and
5'-GACGTGTTAGATCGGAGGCGGCGGCTT(15)-3'. Size selected
1 kb for average insert length 1.9 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine); available through Life
Technologies."

BASE COUNT 185 a 339 c 275 g 159 t
ORIGIN
Alignment Scores: 1.44e-129 Length: 958
Pred. No.: 1114.50 Matches: 226
Score: 94.61% Conservative: 2
Percent Similarity:

Best Local Similarity: 93.788
Query Match: 76.928
DB: 14
Indels: 3
Gaps: 1

US-09-709-103-2 (1-281) x BQ719566 (1-958)

QY 15 SerGluLeuSerIleProAlaLysAsnGlyTyrArgMetValIleLeuGlySerSerLys 34
Db 1 TCGAGAGTGTAGTATCCCGGCAAGAACTGTATCCATGATGATCATCTCGCTCTCCAG 60
QY 35 ValGlyLysThrAlaIleValSerArgPheLeuThrGlyArgPheGluAlaLysThr 54
Db 61 GTGGGCAAGACGGCCATCTGCGGCTTCCACAGGCCCTTCGAGAGCGCTTACAGC 120
QY 55 ProThrIleGluAspPheHisArgLysPheTyrSerIleArgGlyValAlaTyrGlnLeu 74
Db 121 CCGACATCGAGACTTCCACCGCAAGTTCATCCATCCCGCGGAGGTCTACAGCTC 180
QY 75 AspIleLeuAspThrSerGlyAsnHisProPheProAlaMetArgArgLeuSerIleLeu 94
Db 181 GACATCTCTGACACGTCCGCGCAACACCGTTCCTCCCGCCATCGGCGCTCTCATCTC 240
QY 95 ThrGlyAspValPheIleLeuValPheSerLeuAspAsnArgAspSerPheGluVal 114
Db 241 ACAGGAGACGTTTTCATCTGCTGTTCAGTCTGACACACCGCATCTTCGAGAGGTG 300
QY 115 GlnArgLeuArgGlnGlnIleLeuAspThrLysSerCysLeuLysValGlu 134
Db 301 CAGCGGCTCAGGACAGATCTCTGACACCAAGTCTTGCTCAAGAAACCAAGAGAG 360
QY 135 AsnValAspValProLeuValIleCysGlyAsnLysGlyAspArgAspPheTyrArg 154
Db 361 AACGTGAGACGTCCTCGGTGATCTGCGGCAACAGGAGTACCGGACTTACCGGAG 420
QY 155 ValAspGlnArgGlnIleGlnGlnLeuValGlyAspAspProGlnArgCysAlaTyrPhe 174
Db 421 GTGAGCCAGCGCGCATCGACAGCTGTGGCGGAGACCCCGCGCTGCGCTACTTC 480
QY 175 GluIleSerAlaLysLysAsnSerSerLeuAspGlnMetPheArgAlaLeuPheAlaMet 194
Db 481 GAGATCTCGGCGCAAGAAACAGACAGCTGACAGATGTTCCGCGCTCTTCCCATG 540
QY 195 AlaLysLeuProSerGlnMetSerProAspPheLysArgLysValSerValGlnTyrCys 214
Db 541 GCCAAGCTGCCAGAGATGAGCCAGACCTGCCCAAGGCTCTGCGTCACTAATGC 600
QY 215 AspValLeuHisLysLysAlaLeuArgAsnLysLysLeuLeuArg-AlaGlySerGly-G 234
Db 601 GAGCTCTGACACAGAGGCGCTGCGGAACAAAGCTGCGGGCGCGGCAACGAGAG 660
QY 234 LysGlyLysGlyAspPro--GlyAspAlaPheGlyIleValAlaProPheAlaArgArg 252
Db 661 GCCGCGGCGCTGACCCCGGGGGAAGCCTTTTGGGATCGGGGCGCCCTTTCGCGCGC 719

RESULT 7
BM921737 1053 bp mRNA linear EST 12-MAR-2002
LOCUS BM921737
DEFINITION AGENCOURT_6708101 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5753301
5', mRNA sequence.
ACCESSION BM921737
VERSION BM921737.1 GI:19372116
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 1053)
AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.

CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
<http://image.llnl.gov>
Plate: LLM12788 row: n column: 22
High quality sequence stop: 600.
Location/Qualifiers
1..1053
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5753301"
/clone_lib="NIH_MGC_115"
/lab_host="DH10B"
/note="Organ: pooled brain, lung, testis; Vector:
pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC Library."

BASE COUNT 215 a 390 c 297 g 150 t 1 others
ORIGIN

Alignment Scores:

Pred. No.: 9,75e-128 Length: 1053
Score: 1100.50 Matches: 228
Percent Similarity: 88.26% Conservative: 5
Best Local Similarity: 85.36% Mismatches: 22
Query Match: 75.95% Indels: 9
DB: 14 Gaps: 3

US-09-709-103-2 (1-281) x BM921737 (1-1053)

QY 1 MetLysLeuAlaAlaMetIleLysLysMetCysProSerAspSerGluLeuSerIlePro 20
Db 200 ATGAACAGTGGCGGCGGATGATCAGAAAGATGTGCCAGAGACTCGGAGCTGATCCCG 259
QY 21 AlaLysAsnGlyTyrArgMetValIleLeuGlySerSerLysValGlyLysThrAlaLe 40
Db 260 GCCAAGACCTGCTATCGCATGTCATCTCGGCTGCTCCAAAGTGGCAAGAGCGCATC 319
QY 41 ValSerArgPheLeuThrGlyArgPheGluAspAlaTyrThrProThrIleGluAspPhe 60
Db 320 GTGTGCGGCTTCTCTACCGCGCGCTTTCGAGAGCGCTACACCGCTTACCATGAGACTTC 379
QY 61 HisArgLysPheTyrSerIleArgGlyValAlaTyrGlnLeuAspPheIleAspThrSer 80
Db 380 CACCGCAAGTTCATCTCCATCCGCGGCGAGGCTTCACAGCTCGACATCTCGACACGCTCC 439
QY 81 GlyAsnHisProPheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIle 100
Db 440 GGCAACACCCGCTTCCCGCGCATGCGCGCTCTTCATCTCACAGAGACGTTTCATC 499
QY 101 LeuValPheSerLeuAspAsnArgAspSerPheGluValGlnArgLeuArgGlnGln 120
Db 500 CTGGTTCAGTCTGACAAACCGGACATCTTCAGAGAGGTGACGCGGCTCAGCGAGAG 559
QY 121 IleLeuAspThrLysSerCysLeuLysAsnLysThrLysGluAsnValAspValProLeu 140
Db 560 ATCTCTGACACCAAGTCTTCTGCTCAAGAAACCAAGAGAGAGCGTTCCTGATC 619
QY 141 ValIleCysGlyAsnLysGlyAspArgAspPheTyrArgLysValAspGlnArgGluLe 160
Db 620 GTCATCTGCGGCAACAGGAGTACCGGACTTCTACCGGAGGTGACCAAGCGGAGATC 679
QY 161 GlnGlnLeuValGlyAspAspProGlnArgCysAlaTyrPheGluIleSerAlaLysLys 180
Db 680 GAGCAGCTGTGGCGAGACACCCCAAGCGCTGCGCTACTGTGAGATCTCGGCGCAAGAG 739

QY	181	AsnSerSer-LeuAspGlnMetPheArgAlaLeuPheAlaMetLAlaLysLeuProSerG	200
Db	740	AACAGCACCCCTGGAGCCAGATGTCCTCCGGCCCTCTCGCATTGGCCCAAGCGCCACAGGA	799
QY	200	uMetSerProAspLeuHisArgLys---ValSerValGlnTyrCysAspValLeuHisLys	219
		:::	
Db	800	GATGAGCCACGACCTGCACCCCAAGGGGTCTCGGGGCCAAGTACTGCCGCTCGGTGCACAA	859
QY	219	SLYS---AlaLeuArgAsnLysLysLeuLeuArg-AlaGlySer-----G	233
Db	860	AAAGCGGCGCTGGCGGCAACAAAAAACTCTCTCGGGGCGGTAACCCGTTGGCGGGCGCG	919
QY	233	LYGLYGLYGLYGLYAspProGlyAspAlaPheGlyTLeValAlaLProPheAlaArgArgP	253
		:::	:::
Db	920	GAGGAAACCGGGGCGGAACCCCTTTGGGAAATCTGGGGGGAACCCCTTCGTGGGGGCGGGGAC	979
QY	253	roSerVal	255
Db	980	CAAGGGTA	987

RESULT	B
LOCUS	BG085090
DEFINITION	BG108E05-5 N1A Mouse 15K cDNA Clone Set Mus musculus cDNA clone H3108E05 5', mRNA sequence.
ACCESSION	BG085090
VERSION	BG085090.1 GI:1267654
KEYWORDS	EST.
SOURCE	house mouse. Mus musculus
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 649) Karagül,G.J., Dudekula,D.B., Qian,Y., Lim,M.K., Jaradat,S.A., Tanaka .T.S., Carter,M.G. and Ko,M.S.H. Verification and initial annotation of N1A mouse 15K cDNA clone set unpublished (2001)
TITLE	Other-ESTs: H3108E05-3
JOURNAL	Contact: George J. Karagül
COMMENT	Laboratory of Genetics National Institute on Aging/National Institutes of Health 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA Email: conalgsun.grc.nia.nih.gov This clone set has been freely distributed to the community. Please visit http://lgsun.grc.nia.nih.gov/cDNA/15K.html for details. Plate: H3108 row: E column: 05 Seq primer: -21M13 Reverse High quality sequence stop: 649 POLY(A)-No.

```

FEATURES
    source
        Location/Qualifiers
            1. 649
                /organism="Mus musculus"
                /strain="C57BL/6J"
                /db_xref="nia:ST:H3108E05-5"
                /db_xref="taxon:10090"
                /clone="H3108E05"
                /clone_1ID="NTA Mouse 15k cDNA Clone Set"
                /sex="Clones arrayed from a variety of cDNA libraries"
                /dev_stage="Clones arrayed from a variety of cDNA libraries"
                /lab_host="DH10B"
                /note="Vector: pSPOR1; Site.1: SalI; Site.2: NotI; This clone is among a rearranged set of 15,247 clones from 11 embryo cDNA libraries (including preimplantation stage embryos from unfertilized egg to blastocyst, embryonic part of E7.5 embryos, extraembryonic part of E7.5 embryos , and E12.5 female mesonephros/gonad) and one newborn ovary cDNA library. Average insert size 1.5 kb. All source libraries are cloned unidirectionally with Oligo(dT) )-Not primers. References include: (1) Genome-wide expression profiling of mid-gestation placenta and embryo/using a 15,000 mouse developmental cDNA microarray, 2000,

```

BASE COUNT	157 a	182 c	177 g	133 t
ORIGIN				
Alignment Scores:				
pred. No.:	1.85e-123	Length:	649	
score:	1064.00	Matches:	212	
Percent Similarity:	98.62%	Conservative:	2	
Best Local Similarity:	97.70%	Mismatches:	1	
Query Match:	73.43%	Indels:	2	
DB:	12	Gaps:	0	

US-09-709-103-2 (1-281) x BG085090 (1-649)

Oy	18	Set1leProAlaLysAsnGlyTyrArgMetVal1lleuLysSerLysValGlyLys	37
Db	3	AGTATCCGGCCCAAAACATGCTTACAGAGTGGTCTATCTCCGCTCATCCAAAGTGGCCAAAG	62
Oy	38	ThrAla1leValSerArgPheLeuThrGlyArgPheGluAspAlaTyrThrProThr1le	57
Db	63	ACGGCCATTGTGTGGCGTCTCTACAGGGCCGTTTCGAGATGCTTACACCCCTACCATC	122
Oy	58	GluAspPheHisArgLysPheTyrSer1lleArgGlyGlyuAlaTyrGlnLeuAsp1leLeu	77
Db	123	GAGGACCTTCCACCGAAAGTTTATCTCGATCCGGCCGAGCTTACCAAGTTGGACATCTG	182
Oy	78	AspThrSerGlyAsnHisProPheProAlaMetArgAlaGluSer1leLeuThrGlyAsp	97
Db	183	GACATATCCGGCAATCATCCGTTTCCGCCCATGGCCGCCCTCTATCCCTACAGAGAC	242
Oy	98	ValPhe1leLeuValPheSerLeuAspAsnArgAspSerPheGluGluValGlnArgLeu	117
Db	243	GTTTCATTCTGGTGTGTTACGCTTAGACAACCCCGACTATTCGAAAGAGTGCCAAAGCTC	302
Oy	118	ArgGlnGln1leLeuAspPheThrLysSerGlyLeuLysAsnLysThrLysGlnAsnValAsp	137
Db	303	AAACACAGATCTTAAACACCAAGTCTCTCTCMAAGAACMAAACCAAGAGATGTGGAC	362
Oy	138	ValProLeuVal1leCysGlyAsnLysGlyAspArgAspPheTyrArgGluValAspGln	157
Db	363	GTCGCGGTGTCATTTGGCGTTACCAAGGGGAGCCGGGACTTCTTACCGGGAAGTACAGAG	422
Oy	158	ArgGlu1leGluGlnLeuValGlyAspAspProGlnArgCysAlaTyrPheGlu1leSer	177
Db	423	CGGAGAGTTAGAGAGCTGGGGGTGACACCTTCAGCGCTTGTCCTACTCTGAGATCTCA	482
Oy	178	AlaLysLysAsnSerSerLeuAspGlnMetPheArgAla1leuPheAlaMetAlaLysLeu	197
Db	483	GCCAAAGAAAGACAGCAGCTTGACCCAGATGTTCCGTGGCCCTTGGCCATGGCCAAAGCTG	542
Oy	198	ProSerGluMetSerProAspPheLysAsnGlyLysValSerValGlnTyrCys-AspVal1	217
Db	543	CCTGACGAGATGAGACCCCGCACTTGACCCGGAAGGGTATCTGTGCACATCTGGCGACGTAC	602
Oy	217	eunHisLysLysAla1leuArgAsnLysLysLeuLeuArgAlaGlySer	232
Db	603	TGCACAGAAAGCCTCTGAGAACAAAGAAAGCTTCTGGCTGGCGGACAG	649
RESULT 9			
LOCUS	BMS43630	1023 bp	mRNA
DEFINITION	AGNCDCURF_6492527 NIH_MGC_124 Homo sapiens cDNA IMAGE:576764		linear
ACCESSION	BMS43630		
VERSION	BMS43630.1	GI:18774186	
KEYWORDS	EST.		

SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Invitrogen
CDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
Plate: LLM12719 row: m column: 05
High quality sequence stop: 637.

FEATURES

source

1.1023

Location/Qualifiers

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:575214"
/clone_lib="NIH_MGC_124"
/tissue_type="hippocampus"
/lab_host="DH10B"
/note="Organ: brain; Vector: PCMV-SPORT6; Site_1: EcorV (destroyed); Site_2: NotI; RNA source male hippocampus, age 27. Library is oligo-dT primed and directionally cloned (EcorV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 0.9-4 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research genetics tracking code 012."

BASE COUNT

196 a 378 c 281 g 168 t

ORIGIN

Alignment Scores:

Pred. No.: 7.13e-120 Length: 1023
Score: 1038.00 Matches: 222
Percent Similarity: 83.46% Conserved: 5
Best Local Similarity: 81.62% Mismatches: 30
Query Match: 71.64% Indels: 15
DB: 13 Gaps: 3

US-09-709-103-2 (1-281) x BM543630 (1-1023)

QY 1 MettysleuAlaMetilelylMetCysProSerAspSerGluLeuSerilePro 20
Db 149 ATGAAGCTGCGCCGATGATCAAGAGATGTGCCCGACGACATCGAGTATCCCG 208
QY 21 AlalysancysTyrArgetvalileleuGlySerSerlyValGlylysthrAlaile 40
Db 209 GCCAAGAACTGCTATCGATGTCATCTCGGCTCTCCCAAGTGGGCAAGGGCCATC 268
QY 41 ValSerArgheleumrgrlyArgpneGluAspAlaTyrThrProThrileGluAspHe 60
Db 269 GTGTGGCTTCCCTCACCGCGCTTCGAGAGCCCTACACGCGCTACATCGAGGACCTTC 328
QY 61 HisArglySerPheTyrSerileArgGlyGluValTyrGlnLeuAspIleLeuAspThrSer 80
Db 329 CACCGCAAGTTTACATCCATCCCGCGAGGTCTACCACTGACATCTCCGACACGCTCC 388
QY 81 GlyAsnHisProPheProAlaMetArgArgLeuSerileLeuThrGlyAspValPheile 100
Db 389 GGCACCAACCCGTTCCCGCATGCGCGCTCTCCATCTCCACAGAGAGCTTTTATC 448
QY 101 LeuValPheSerleuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGlnGln 120
Db 449 CTGGTTCAGTCTGACACACCGGCACTCTTCGAGAGGTGACACGCGCTCAGGACGAG 508
QY 121 IleLeuAspThrlySerCysLeuAsnlysthrlyGluAsnValAspValProleu 140

Db 509 ATCTCCACACCAAGTCTGCTCAAGAACAAACACAGAGAGACTGACTGCCCTG 568
QY 141 ValTleCysGlyAsnlysthrlyAspArgAspPheTyrArgGluValAspGlnArgGluile 160
Db 569 GTATCTGCGGCAACAGAGGTGACCGCTTCTTACCGGAGGTGACGAGCGAGATC 628
QY 161 GluGlnLeuValGlyAspAspProGlnArgCysAlaTyrPheGlnIleSerAlaIlys 180
Db 629 GAGAGGTGCTGGGACACGACCCCGAGCGCTGCTTCTGAGATCTCCGACAGAG 688
QY 181 AsnSerSerleu-AspGlnMetPheArgAlaLeuPheAlaMetAla-LysLeu-ProSer 199
Db 689 AACACAGCTTGGGACCAAGTGTCTCGCGCTTCTCCATGGCCCAAGTGCCTCCGCG 748
QY 200 GluMetSerProAspPheHisArgly--ValSerValGlnTyrCys--AspValLeu 217
Db 749 GAGATGAGCCCAAGCTGACCGCCCAAGGTCTCGGCTGAGGTGAGTGGCAGCGGCTGG 808
QY 218 HislyleuAlaLeuArgAsnlysthrlyleuArgAlaGlySerGlyGlyGly 237
Db 809 CACAGAAAGCGCGTGGCGGACACAGAAAGTGTCTCGGCGCGGCGGCAATCGCGGCG 868
QY 238 -----AspProGlyAspAlaPheGlyIleValAla 247
Db 869 CCCTCTCTAGTCATTAACCTTGACCGAAGGACCTTGGGCGAGCTTGGGCGCACCATTTT 928
QY 248 PropheAlaArgArgProSerValHisSer 257
Db 929 CCGGACAGTGGCGCTCTGTGTGTCATCC 958

RESULT 10
B0067637
LOCUS B0067637
DEFINITION AGENCOURT_6759053 NIH_MGC_115 Homo sapiens CDNA clone IMAGE:575214
ACCESSION B0067637
VERSION B0067637.1 GI:19896683
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
Plate: LLM12793 row: n column: 15
High quality sequence stop: 602.

FEATURES

source

1.1032

Location/Qualifiers

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:575214"
/clone_lib="NIH_MGC_115"
/lab_host="DH10B"
/note="Organ: pooled brain, lung, testis; Vector: PCMV-SPORT6; Site_1: NotI; Site_2: EcorV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27; 1 oligo-dT primed and directionally cloned (EcorV site is destroyed upon cloning). Average insert size 1.8 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C.

QY 103 PheserleuspsasnaatgaspserPhegluGlulValGlnArgleuArgGlnGlnleu 122
 DB 355 TTGAGCCTGGACACAGGAGCTCTTTGAGAGAGTGCAGAGCTGAAGCAGCATATG 414
 QY 123 AsphrlysserCysleuLysAsnLysThrLysGlnAsnValAspValProleuValIle 142
 DB 415 GAGACCAAGCTCTCTTCAAGAACACCAAGAGAGATGTTCCCATCGTTATC 474
 QY 143 CysGlyAsnLysGlyAspArgAspPheTyrArgIuValAspGlnArgIuGlnGln 162
 DB 475 TGGCGGAATTAAGTACAGAGAGCTTACAGGAAGTGCAGCTCATGATGACAG 534
 QY 163 LeuValGlyAspAspProGlnArgCysAlaTyrPheGluIleSerAlaLysLysAsnSer 182
 DB 535 CTGGTTGGAAGAGGAC--AGTAAATGCTCTTACTTGAAGGTGCGGCAAGAAACTC 591
 QY 183 SerleuAspGlnMetPheArgAlaLeuPheAlaMetAlaLysLeuProserGlnMetSer 202
 DB 592 AGCCTGGATGAGATGTTAAAGCTCTTCCATCATGACAGATGCGCCAGCGAGATGAGC 651
 QY 203 ProAspLeuHisArgLysValSerValGlnTyrCysAspValLeuHisLysLysAlaLeu 222
 DB 652 TCAGACTGACCGCAGAGGTGTCTGCAGTACTGTGAGATTTTACACAAAGAGTCCCTG 711
 QY 223 ArgAsnLysLysLeu 227
 DB 712 AAGAAAGAAAGTA 726

RESULT 12
 BG711792 655 bp mRNA linear EST 08-MAY-2001
 LOCUS BG711792
 DEFINITION pgl1n.pk009.115 Normalized Liver Library Gallus gallus cDNA clone
 pgl1n.pk009.115 5' similar to gb|AA03090.1|AF239157.1 (AF239157)
 DEKRA1 (Rattus norvegicus)6, mRNA sequence.
 BG711792
 ACCESSION BG711792.1 GI:14005742
 VERSION EST.
 KEYWORDS chicken.
 SOURCE Gallus gallus
 ORGANISM Gallus gallus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianidae; Gallus.
 REFERENCE 1 (bases 1 to 655)
 AUTHORS Burnside,J., Morgan,R.W. and Coopburn,L.A.
 TITLE Chicken ESTs from a normalized liver library
 JOURNAL Unpublished (2001)
 COMMENT Contact: Joan Burnside
 Molecular Endocrinology
 University of Delaware
 40 Townsend Hall, Newark, DE 19717, USA
 Tel: 302 831-1345
 Fax: 302 831-3411
 Email: joan@udel.edu, www.chickest.udel.edu.

FEATURES
 source
 1. .655
 Location/Qualifiers
 /organism="Gallus gallus"
 /db_xref="taxon:9031"
 /clone_lib="Normalized Liver Library"
 /sex="Male and Female"
 /tissue_type="liver"
 /lab_host="E.coli EMDH10B"
 /note="Vector: pCMVSPORT 6"

BASE COUNT 145 a 206 c 184 g 104 t 16 others
 ORIGIN

Alignment Scores:
 Pred. No.: 9.68e-114 Length: 655
 Score: 987.00 Matches: 192
 Percent Similarity: 97.12% Conservative: 10
 Best Local Similarity: 92.31% Mismatches: 5
 Query Match: 68.12% Indels: 1

DB: 12 Gaps: 0
 US-09-709-103-2 (1-281) x BG711792 (1-655)

QY 1 MetLysLeuAlaAlaMetIleLysLysMetCysProSerAspSerGlnLeuSerIlePro 20
 DB 14 AAGAACTGCGCAGCATGATCAAGAAAGATGTGTCCACAGAGCGTACAGATCCCC 73
 QY 21 AlaLysAsnCysTyrArgMetValIleLeuGlySerSerLysValGlyLysThrAlaIle 40
 DB 74 GCCAAGAACTGCTACCGCATGTGTCTCTGCTCTCCAAAGGTGGCAAGCGCCATC 133
 QY 41 ValSerArgPheLeuThrGlyArgPheGluAspAlaTyrTrpProThrIleGluAspPhe 60
 DB 134 GTCTGGCGGTTCTTACCGCGCGCTTCAGAGAGAGTACAGCCCATGAGAGACTTC 193
 QY 61 HisArgLysPheTyrSerIleArgGlyGlyValTyrGlnLeuAspIleLeuAspThrSer 80
 DB 194 CACCCCAAGTTCTACAGCATCCGCGGTGAGGTCTACAGCTCGACATCTCGACACGCTG 253
 QY 81 GlyAsnHisProPheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIle 100
 DB 254 GGCAACCACTCCCTCCCGCATGCGCGCTGTCTCATCTTCACAGAGTACAGGTTTTCATC 313
 QY 101 LeuValPheSerLeuAspAsnArgAspSerPheGluGlnValGlnArgLeuArgGln 120
 DB 314 CTCTGTTACAGCTTGAGACACCGGAGCTCTTCAGAGAGTGCAGCGCTTGACACAGCAG 373
 QY 121 IleLeuAspThrLysSerCysLeuLysAsnLysThrLysGlnAsnValAspValProleu 140
 DB 374 ATCTGTGAGACCAAGTCTCTGCTCTCAAGAACAAACCAAGAGACATCAGAGTCCCGCTG 433
 QY 141 ValIleCysGlyAsnLysGlyAspArgAspPheTyrArgIuValAspGlnArgIuIle 160
 DB 434 GTCACTCTCGGCAACCAAGGCGAGCGGACTTTTACGGGGGGGTGACGCCCGAGAGATC 493
 QY 161 GluGlnLeuValGlyAspAspProGlnArgCysAla-TyrPheGluIleSerAlaLysLys 180
 DB 494 GAGCAGCTGTGGCGCAGACCCCAAGAAATGCCCNNACTTCAGATCTCGGCGCAAGAA 533
 QY 180 SASerSerleuAspGlnMetPheArgAlaLeuPheAlaMetAlaLysLeuProserGln 200
 DB 554 GACACGACGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 613
 QY 200 uMetSerProAspLeuHisArg 207
 DB 614 GATGAGCCCGGACCTGCACCGC 635

RESULT 13
 B1553776 758 bp mRNA linear EST 05-SEP-2001
 LOCUS B1553776
 DEFINITION 603190722F1 NIH_MGC_95 Homo sapiens cDNA clone IMAGE:5262101 5',
 mRNA sequence.
 B1553776
 ACCESSION B1553776
 VERSION B1553776.1 GI:15441088
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 758)
 AUTHORS NIH-MGC http://mgc.ncl.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cyabbs-remail.nih.gov
 Tissue Procurement: Miklos Palcovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiroki
 Toshiyuki and Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:

US-09-709-103-2 (1-281) x B1553776 (1-758)

Db 546 CTGGCTCATCAGC 558

source

1. .1137

BASE COUNT	223 a	427 c	315 g	172 t
ORIGIN				

DB:	14	Gaps:	3
-----	----	-------	---

US-09-/09-103-2 (1-281) x BM921656 (1-1137)

Db 401 CACCGCAAGTTCCTACTCCATCCGCGCGCAGAGTCTACCGAGCTCGACACATTCCTCGACACGTC 460

```

QY      81 glyAsnHisProPheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIle 100
      461 GGCACACACCCCTCCGCCCATCCGCGGCCCTCCATCCATCCACAGAGAGCGTTTTCATC 520
QY      101 LeuValPheSerLeuAspAsnArgAspSerPheGluGluValAlaArgLeuArgGlnGln 120
      521 CTGGGTCTAGTCTGGACAAACCGCACATCTCTCGAGGAGTCCACGGCTCAGCGACGAG 580
QY      121 IleLeuAspThrLysSerCysLeuLysAsnLysThrLysGluAsnValAspValProLeu 140
      581 ATCTCTGCACCAAGCTTGCTCCCAAGAACAAACCAAGAGATGTGACGTGCGCCCTG 640
QY      141 ValIleCysGlyAsnLysGlyAspArgAspPheTyrArgGluVal-AspGlnArgGluIle 160
      641 GTCATCTGGCGGCAACACAGATGACCGGCTCTACCGCGAGGTGGAGACACCGGAGAT 700
QY      160 eglGluGlnLeuValAlaGlyAspArgProGlnArgCysAlaTyrPheGluIleSerAlaLys 180
      701 CGAGCAGCTGTGTGGGCGG-ACACCCACGCGGTGCGCTTCTCGAGATCTGCGCCAGAA 759
QY      180 AsnSer-SerLeuAsp-GlnMetPheArgAlaLeuPheAlaMetAlaLys-LeuProse 199
      760 AACACATATCCCTGACCCCAATGTCTCCGCGCTCTCCCTCCCTCCACACCTGCGCCAG 819
QY      199 T-----GluMetSerProAspLeuHisArgLysValSerValGlnTyrCys 214
      820 CGGAGACAGCAGTCCCGGACCTTGACCCCAAGTCTCCGCGGAGGAGATACGCGCAACC 879
QY      214 sAspValIleuHisLysAlaLeuArgAsnLysLysLeuLeuArg-AlaGlySerGlyG 234
      880 GGGTGGCCCATAAAGGGGCGCTGCCGAAACAAAGAAATGCTGCCGGGAGCGGCC 939
QY      234 LysLysGlyLysAspPro----- 239
      940 ACCGCGCTGGGTGGCCGTCTGTCCACCGCGGGGGAACGTCTCTGTGTGGCGAACC 999
QY      240 --GlyAsp-----AlaPheGlyIleValAlaProPheAlaArgProSerValH 256
      1000 GGGGGAGACCCCTTGTCGGCGGGCGGTGCGACCAATGTGTACAAACCGGCCAACCCCTC 1059
QY      256 IsSerAspLeuMetTyrIleArgGlnLysAlaSer 267
      1060 CCCGGTTACACTCCCTCCGCCAATAAAGGCTC 1094

RESULT 15
LOCUS   B1596688 814 bp mRNA linear EST 07-SNP-2001
DEFINITION
603243262F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5285894 5',
mRNA sequence.
ACCESSION B1596688
VERSION B1596688
KEYWORDS B1596688.1 GI:15489627
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 814)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLES National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: egaups-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shitaki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LIML at:
http://image.llnl.gov
Plate: L1AM11722 row: k column: 15
High quality sequence stop: 803.
Location/Qualifiers

```

```

source
1. .814
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5285894"
/clone_11b="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/Note="Organ: brain; Vector: pBluescript (modified
pBluescript KS+); Site:1: BamHI; Site:2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-tttttttttttttttVN-3',
size-selected for average insert size 2.3 kb and
normalized to R0T 5. This is a primary library enriched
for full-length clones and constructed using the
cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC library."

BASE COUNT 160 a 310 c 225 g 119 t
ORIGIN

Alignment Scores:
Pred. No.: 9.15e-105 Length: 814
Score: 917.00 Matches: 192
Percent Similarity: 96.48% Conservative: 0
Best-Local Similarity: 96.48% Mismatches: 5
Query Match: 63.29% Indels: 4
DB: 13 Gaps: 0

US-09-709-103-2 (1-261) x B1596688 (1-814)
QY      1 MetLysIleuAlaAlaMetIleLysLysMetCysProSerAspSerGluLeuSerIlePro 20
      220 ATGAATAGTGGCGCGCATATATCAAGAAATGTGTGCCGAGACGACTGGAGTGAATATCCG 279
QY      21 AlaLysAsnCysTyrArgMetValIleLeuGlySerSerLysValGlyLysThrAlaIle 40
      280 GCCAAGAACTGCTATTCGATGTGTCATCTCGGCTGTCTCAAGTGGGCAAGACGGCCATC 339
QY      41 ValSerArgPheLeuThrGlyArgPheGluAspArgIleThrProThrIleGluAspPhe 60
      340 GTGTGCGCGCTTCTCAGCGCGCTTCAGAGACGCTCAACGCTTACCATGCAAGAGACTTC 399
QY      61 HisArgLysPheTyrSerIleArgGlyGluValTyrGlnLeuAspIleLeuAspThrSer 80
      400 CACCGAAATTTACTCATTCATCGCGGCGAGGTCTACAGCTGACATCTCGACACGTCC 459
QY      81 GlyAsnHisProPheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIle 100
      460 GGCACACACCCCTCCGCCCATCCGCGGCCCTCCATCCATCCACAGAGAGCGTTTTCATC 519
QY      101 LeuValPheSerLeuAspAsnArgAspSerPheGluGluValAlaArgLeuArgGlnGln 120
      520 CTGGGTCTAGTCTGGACAAACCGCACATCTCTCGAGGAGTCCACGGCTCAGCGACGAG 579
QY      121 IleLeuAspThrLysSerCysLeuLysAsnLysThrLysGluAsnValAspValProLeu 140
      580 ATCTCTGCACCAAGCTTGCTCCCAAGAACAAACCAAGAGATGTGACGTGCGCCCTG 639
QY      141 ValIleCysGlyAsnLysGlyAspArgAspPheTyrArgGluValAspGlnArgGluIle 160
      640 GTCATCTGGCGGCAACACAGATGACCGGCTCTACCGCGAGGTGGAGACACCGGAGAT 699
QY      161 GluGlnLeuValAlaGlyAspArg-ProGlnArgCysAlaTyrPheGlu-IleSerAlaLys 180
      700 GACACACTGTGTGGGCGGAGACCCCGAGCGCTGCGCTTCTCGAAGATCTGCGCCAGAA 759
QY      180 yAsnSerSerLeuAspGlnMetPheArgAlaLeuPheAlaMetAlaLysLeu 197
      760 AGAACAGACAGCCG-GACCAAGATGC-CGGCGCTCTTCCCATGAGCAAGCTG 810

Search completed: December 28, 2002, 02:47:21
Job time : 2197 secs

```

61 GCCAAGACTGCTATCGCATGTGTCATCCTCGGCTCCTCCAAAGTGGGCAAGAGCGCCATC 120

```

Db 315 GCCAAGACTGTATGATGATGATCTGCTGCTCCAGAGTGGGCAAGACGGCCATC 374
QY 121 GTGTGGCGCTTCTCTCAACGGCGGCTTGGAGACGCTTACACGGCTTACATCGAGACTTC 180
Db 375 GTGTGGCGCTTCTCTCAACGGCGGCTTGGAGACGCTTACACGGCTTACATCGAGACTTC 434
QY 181 CACCGCAAGTTCTTACATCTCCGCGGCGAGGTTACACGCTTACATCTCTCGACACGTC 240
Db 435 CACCGCAAGTTCTTACATCTCCGCGGCGAGGTTACACGCTTACATCTCTCGACACGTC 494
QY 241 GGCACACACCGCTTCCCGCGGAGGCGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 300
Db 495 GGCACACACCGCTTCCCGCGGAGGCGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 554
QY 301 CTGGTGTTCAGTCTGGACACCGGAGCTCTTCTTGAAGAGTGGAGGCTTCTCTCTCTCTCT 360
Db 555 CTGGTGTTCAGTCTGGACACCGGAGCTCTTCTTGAAGAGTGGAGGCTTCTCTCTCTCTCT 614
QY 361 ATCTCGACACCAAGTCTTCTCTCAAGAACCAAGAGAGAGCTTGAAGCTTGGCCCTTG 420
Db 615 ATCTCGACACCAAGTCTTCTCTCAAGAACCAAGAGAGAGCTTGAAGCTTGGCCCTTG 674
QY 421 GTATCTGCGGCAACAGGTTGACCGGAGTCTTACCGGCAAGTGGACGCGCGGAGATC 480
Db 675 GTATCTGCGGCAACAGGTTGACCGGAGTCTTACCGGCAAGTGGACGCGCGGAGATC 734
QY 481 GAGCAGCTGGTGGGCGACGACCGCCAGCGCTTGGCGCTTCTCTCTCTCTCTCTCTCTCTCT 540
Db 735 GAGCAGCTGGTGGGCGACGACCGCCAGCGCTTGGCGCTTCTCTCTCTCTCTCTCTCTCTCT 794
QY 541 AACGACAGCTTGGACACAGATGTTCCGCGGCTTGGCGCTTGGCGCTTGGCGCTTGGCGCTT 600
Db 795 AACGACAGCTTGGACACAGATGTTCCGCGGCTTGGCGCTTGGCGCTTGGCGCTTGGCGCTT 854
QY 601 ATGAGCCCAAGCTTGGACACAGATGTTCCGCGGCTTGGCGCTTGGCGCTTGGCGCTTGGCG 660
Db 855 ATGAGCCCAAGCTTGGACACAGATGTTCCGCGGCTTGGCGCTTGGCGCTTGGCGCTTGGCG 914
QY 661 GCGCTGGGGAACAAGAGCTTGTGCGGCGCGGACGCGCGGCGGCGGCGGCGGCGGCGGCGG 720
Db 915 GCGCTGGGGAACAAGAGCTTGTGCGGCGCGGACGCGCGGCGGCGGCGGCGGCGGCGGCGG 974
QY 721 GAGCGCTTGGCAATCTGTGCGACCTTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 780
Db 975 GAGCGCTTGGCAATCTGTGCGACCTTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1034
QY 781 TACATCCGCGAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 840
Db 1035 TACATCCGCGAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1094
QY 841 AGCTAG 846
Db 1095 AGCTAG 1100

```

RESULT 2
US-09-053-374A-4
Sequence 4, Application US/09053374A
Patent No. 6462177

GENERAL INFORMATION:

APPLICANT: YEN KWANG-MU
TITLE OF INVENTION: MAMMALIAN BLOOD LOSS-INDUCED GENE, KD312
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: AMGEN INC.
STREET: ONE AMGEN CENTER DRIVE
CITY: THOUSAND OAKS
STATE: CA
COUNTRY: US
ZIP: 91320
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

```

? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: PatentIn Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/09/053,374A
? FILING DATE:
? CLASSIFICATION: 435
? ATTORNEY/AGENT INFORMATION:
? NAME: COOK, ROBERT R.
? REGISTRATION NUMBER: 31,602
? REFERENCE/DOCKET NUMBER: A-514
? INFORMATION FOR SEQ ID NO: 4:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1689 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: cDNA
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 132..971
? US-09-053-374A-4

```

Query Match

Best Local Similarity 80.1%; Score 678; DB 4; Length 1689;

Matches 748; Conservative 0; Mismatches 95; Indels 3; Gaps 1;

```

QY 1 ATGAAACTGGCGCGATGATCAAGAGATGTCGCGAGCTCGGAGCTGAGTATCCCG 60
Db 132 ATGAAACTGGCGCGATGATCAAGAGATGTCGCGAGCTCGGAGCTGAGTATCCCG 191
QY 61 GCCAAGAACTGCTATCGCATGCTGCTCCGCGCTGCTCAAGGTTGGGCAAGCGGCGATC 120
Db 192 GCCAAGAACTGCTATCGCATGCTGCTCCGCGCTGCTCAAGGTTGGGCAAGCGGCGATC 251
QY 121 GTGTGCGGCTTCTCTCAACGGCGGCTTGGAGACGCTTACACGCTTACACCTTGAAGACTTC 180
Db 252 GTGTGCGGCTTCTCTCAACGGCGGCTTGGAGACGCTTACACGCTTACACCTTGAAGACTTC 311
QY 181 CACCGCAAGTTCTTACATCTCCGCGGCGGAGTCTTACAGCTTGAACATCTCTGACACGTC 240
Db 312 CACCGCAAGTTCTTACATCTCCGCGGCGGAGTCTTACAGCTTGAACATCTCTGACACGTC 371
QY 241 GGCACACACCGCTTCCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
Db 372 GGCACACACCGCTTCCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 431
QY 301 CTGGTGTTCAGTCTGGACACCGGAGCTCTTCTGAGAGCTTCTGAGGCTTCTGAGGCTTCT 360
Db 432 CTGGTGTTCAGTCTGGACACCGGAGCTCTTCTGAGAGCTTCTGAGGCTTCTGAGGCTTCT 491
QY 361 ATCTCGACACCAAGTCTTCTCTCAAGAACCAAGAGAGAGTGGAGCTTGGCCCTTG 420
Db 492 ATCTCGACACCAAGTCTTCTCTCAAGAACCAAGAGAGAGTGGAGCTTGGCCCTTG 551
QY 421 GTCATCTGGGGAACAAGGTTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 480
Db 552 GTCATCTGGGGAACAAGGTTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 611
QY 481 GAGCAGCTGGTGGGCGACGACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 540
Db 612 GAGCAGCTGGTGGGCGACGACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 671
QY 541 AACGACAGCTTGGACACAGATGTTCCGCGGCTTGGCGCTTGGCGCTTGGCGCTTGGCGCTT 600
Db 672 AACGACAGCTTGGACACAGATGTTCCGCGGCTTGGCGCTTGGCGCTTGGCGCTTGGCGCTT 731
QY 601 ATGAGCCCAAGCTTGGACACAGATGTTCCGCGGCTTGGCGCTTGGCGCTTGGCGCTTGGCG 660
Db 732 ATGAGCCCAAGCTTGGACACAGATGTTCCGCGGCTTGGCGCTTGGCGCTTGGCGCTTGGCG 791
QY 661 GCGCTGGGGAACAAGAGCTTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 720
Db 792 GCGCTGGGGAACAAGAGCTTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 848

```

QY 721 GAGCGCTTGGCATCGTGGACACCTTGGCGCGCCGCCACAGCTACAGACGACTCATG 780
|||
Db 849 GATCCCTTGGCATCTTGGCGCCCTTGGCTGCGAGACCTAGCGCTACGACCTCATG 908
QY 781 TACATCCGCGAGAGCG 840
|||
Db 909 TACATCTGTGAGAAACAGCTGTGAGCGACGCGCTAGAGCGAGCGCGCTGTGTATC 968
QY 841 AGCTAG 846
|||
Db 969 AGTAG 974

RESULT 3
US-09-053-374A-3
; Sequence 3, Application US/09053374A
; Patent No. 6462177
; GENERAL INFORMATION:
; APPLICANT: YEN, KWANG-MU
; TITLE OF INVENTION: MAMMALIAN BLOOD LOSS-INDUCED GENE, KD312
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: AMGEN INC.
; STREET: ONE AMGEN CENTER DRIVE
; CITY: THOUSAND OAKS
; STATE: CA
; COUNTRY: US
; ZIP: 91320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/053,374A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: COOK, ROBERT R.
; REGISTRATION NUMBER: 31,602
; REFERENCE/DOCKET NUMBER: A-514
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3986 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-09-053-374A-3

Query Match 73.9%; Score 625; DB 4; Length 3986;
Best Local Similarity 80.0%; Pred. No. 1.6e-113;
Matches 846; Conservative 0; Mismatches 0; Indels 211; Gaps 1;

QY 1 ATGAACTGGCGGATGATCAAGAGATGTGCCGAGCGACCTCGAGCTGATATCCG 60
|||
Db 774 ATGAACTGGCGGATGATCAAGAGATGTGCCGAGCGACCTCGAGCTGATATCCG 833
QY 61 GCCAAGAACTGATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120
|||
Db 834 GCCAAGAACTGATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 893
QY 121 GTGTGCGGCTTCTGACGCGCGCTTGAAGAGAGCGCTTACAGCTGAGAGCTTC 180
|||
Db 894 GTGTGCGGCTTCTGACGCGCGCTTGAAGAGAGCGCTTACAGCTGAGAGCTTC 953
QY 181 CACCGCAAGTTTCACTCATCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
|||
Db 954 CACCGCAAGTTTCACTCATCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1013
QY 241 GGCACCAACCGCTTCCCGCATGCGGCGCTCTCATCTCTAC----- 284
|||

Db 1014 GGCACCAACCGCTTCCCGCATGCGGCGCTCTCATCTCTACAGAGTGAAGCGGGGGCC 1073
QY 285 ----- 284
Db 1074 GGCAGAGTGGGAG 1133
QY 285 ----- 284
Db 1134 CCGGCTGTGCGCGCGCGAGTAGTGCCTTTCGCGCTTGAAGAGGTAGAGCGCGCGCG 1193
QY 285 ----- 284
Db 1194 GCCTCAAGTCAAGCCGACTTGTCCCTGGCGCGCCACCCCTACCTTCTCTTCTGCT 1253
QY 285 -----AGGAGAGTCTTTCATCTCTGCTTTCAGTCTGAGCAACCGGAGTCT 329
Db 1254 CTCTGTGCCCCCTTAG 1313
QY 330 CTTCGAGAGAGTGCAGAGGCTCAGGAGAGAGATCTTCGACACCAAGTCTTGCCTAAG 389
Db 1314 CTTCGAGAGAGTGCAGAGGCTCAGGAGAGAGATCTTCGACACCAAGTCTTGCCTAAG 1373
QY 390 CAAAACCAAG 449
Db 1374 CAAAACCAAG 1433
QY 450 CTCTACCGCGAGAGTGCAG 509
Db 1434 CTCTACCGCGAGAGTGCAG 1493
QY 510 CTTCGAGAGAGTGCAG 569
Db 1494 CTTCGAGAGAGTGCAG 1553
QY 570 GCTCTGCGGATGAGGAG 629
Db 1554 GCTCTGCGGATGAGGAG 1613
QY 630 GGTGAGTACTGAG 689
Db 1614 GGTGAGTACTGAG 1673
QY 690 CGGAG 749
Db 1674 CGGAG 1733
QY 750 GCGCGCGCGCGAG 809
Db 1734 GCGCGCGCGCGAG 1793
QY 810 CCAGGCCAAG 846
Db 1794 CCAGGCCAAG 1830

RESULT 4
US-09-053-374A-6
; Sequence 6, Application US/09053374A
; Patent No. 6462177
; GENERAL INFORMATION:
; APPLICANT: YEN, KWANG-MU
; TITLE OF INVENTION: MAMMALIAN BLOOD LOSS-INDUCED GENE, KD312
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: AMGEN INC.
; STREET: ONE AMGEN CENTER DRIVE
; CITY: THOUSAND OAKS
; STATE: CA
; COUNTRY: US
; ZIP: 91320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible


```
;; GENERAL INFORMATION:
;; APPLICANT: Chien, Shu
;; TITLE OF INVENTION: GENE THERAPY IN CORONARY ANGIOPLASTY AND
;; FILE REFERENCE: BYPASS
;; CURRENT APPLICATION NUMBER: US/08/884,866A
;; PRIOR FILING DATE: 1997-06-30
;; PRIOR APPLICATION NUMBER: 60/030,358
;; NUMBER OF SEQ ID NOS: 25
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 9
;; LENGTH: 480
;; TYPE: DNA
;; ORGANISM: Rattus norvegicus
;; NAME/KEY: misc_feature
;; LOCATION: (49)...(51)
;; OTHER INFORMATION: nnn = Any nucleic acid triplet, except for UCA,
;; OTHER INFORMATION: UCC, UCG, UCU, AGC, or AGU
US-08-884-866A-9
```

```
Query Match          10.6%; Score 89.6; DB 4; Length 480;
Best Local Similarity 56.4%; Pred. No. 2e-09;
Matches 164; Conservative 0; Mismatches 127; Indels 0; Gaps 0;
```

```
QY 73 TATCGATGTCATCTCGCTGCTCCAGAGTGGCGACAGCGCCATCTGTCGGCTTC 132
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 10 TATAAGCTGTGTGTGGTGGCGCGGCTGTGGGCAANNNGGCGTACATCCAGCTG 69
QY 133 CTCACCGCGCGCTTCGAGAGCGCTTCACAGCCTTACCATGAGAGCTTCACCGCAATTTC 192
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 70 ATCCAGAACCACTTTTGTGTGAGCAATACGACCCCACTATGAGAAATTCCTACCGGAGCAG 129
QY 193 TACTCCATCCGCGCGAGGTCTACAGCTCGACATCTCGACACGCGCGCAACCAACCGG 252
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 130 GTGTGATTTGATGGGAGAGAGCTGCTGTGTGACATCTGATACCGCGCGCTGGAGAG 189
QY 253 TTCGCCGCGATGGCGGCGCTCTCCATCTTCACAGAGAGAGCTTTTCATCTGTGTTCAGT 312
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 190 TACAGCGCATGGGAGGAGCAATACGACCGGAGGAGGCTTCTGCTGTGTTCGCC 249
QY 313 CTGGAACACCGGAGCTCTTCAGAGAGGTGACAGCGGCTAGGACAGCATC 363
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 250 ATCAACAACCAAGCTTTTGTGAGGACATCCACAGTACAGGAGAGCATC 300
```

RESULT 7

```
;; Sequence 11, Application US/08884866A
;; GENERAL INFORMATION:
;; APPLICANT: Chien, Shu
;; TITLE OF INVENTION: GENE THERAPY IN CORONARY ANGIOPLASTY AND
;; FILE REFERENCE: BYPASS
;; CURRENT APPLICATION NUMBER: US/08/884,866A
;; PRIOR FILING DATE: 1997-06-30
;; PRIOR APPLICATION NUMBER: 60/030,358
;; NUMBER OF SEQ ID NOS: 25
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 11
;; LENGTH: 570
;; TYPE: DNA
;; ORGANISM: Artificial Sequence
;; NAME/KEY: misc_feature
;; LOCATION: (49)...(51)
;; OTHER INFORMATION: nnn = Any nucleic acid triplet, except for UCA,
;; OTHER INFORMATION: UCC, UCG, UCU, AGC, or AGU
;; OTHER INFORMATION: Variation of SEQ ID NO.: 2
```

US-08-884-866A-11

```
Query Match          10.6%; Score 89.6; DB 4; Length 570;
Best Local Similarity 56.4%; Pred. No. 2e-09;
Matches 164; Conservative 0; Mismatches 127; Indels 0; Gaps 0;
```

```
QY 73 TATCGATGTCATCTCGCTGCTCCAGAGTGGCGACAGCGCCATCTGTCGGCTTC 132
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 10 TATAAGCTGTGTGTGGTGGCGCGGCTGTGGGCAANNNGGCGTACATCCAGCTG 69
QY 133 CTCACCGCGCGCTTCGAGAGCGCTTCACAGCCTTACCATGAGAGCTTCACCGCAATTTC 192
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 70 ATCCAGAACCACTTTTGTGTGAGCAATACGACCCCACTATGAGAAATTCCTACCGGAGCAG 129
QY 193 TACTCCATCCGCGCGAGGTCTACAGCTCGACATCTCGACAGCTGCGGCAACCAACCGG 252
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 130 GTGTGATTTGATGGGAGAGAGCTGCTGTGTGACATCTGATACCGCGCGCTGGAGAG 189
QY 253 TTCGCCGCGATGGCGGCGCTCTCCATCTTCACAGAGAGCTTTTCATCTGTGTTCAGT 312
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 190 TACAGCGCATGGGAGGAGCAATACGACCGGAGGAGGCTTCTGTGTGTTCGCC 249
QY 313 CTGGAACACCGGAGCTCTTCAGAGAGGTGACAGCGGCTAGGACAGCATC 363
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 250 ATCAACAACCAAGCTTTTGTGAGGACATCCACAGTACAGGAGAGCATC 300
```

RESULT 8

```
;; Sequence 2, Application US/08884866A
;; GENERAL INFORMATION:
;; APPLICANT: Chien, Shu
;; TITLE OF INVENTION: GENE THERAPY IN CORONARY ANGIOPLASTY AND
;; FILE REFERENCE: BYPASS
;; CURRENT APPLICATION NUMBER: US/08/884,866A
;; PRIOR FILING DATE: 1997-06-30
;; PRIOR APPLICATION NUMBER: 60/030,358
;; NUMBER OF SEQ ID NOS: 25
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 2
;; LENGTH: 570
;; TYPE: DNA
;; ORGANISM: Rattus norvegicus
;; NAME/KEY: CDS
;; LOCATION: (1)...(570)
US-08-884-866A-2
```

```
Query Match          10.6%; Score 89.4; DB 4; Length 570;
Best Local Similarity 56.7%; Pred. No. 2e-09;
Matches 165; Conservative 0; Mismatches 126; Indels 0; Gaps 0;
```

```
QY 73 TATCGATGTCATCTCGCTGCTCCAGAGTGGCGACAGCGCCATCTGTCGGCTTC 132
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 10 TATAAGCTGTGTGTGGTGGCGCGGCTGTGGGCAANNNGGCGTACATCCAGCTG 69
QY 133 CTCACCGCGCGCTTCGAGAGCGCTTCACAGCCTTACCATGAGAGCTTCACCGCAATTTC 192
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 70 ATCCAGAACCACTTTTGTGTGAGCAATACGACCCCACTATGAGAAATTCCTACCGGAGCAG 129
QY 193 TACTCCATCCGCGCGAGGTCTACAGCTCGACATCTCGACAGCTGCGGCAACCAACCGG 252
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 130 GTGTGATTTGATGGGAGAGAGCTGCTGTGTGACATCTGATACCGCGCGCTGGAGAG 189
QY 253 TTCGCCGCGATGGCGGCGCTCTCCATCTTCACAGAGAGCTTTTCATCTGTGTTCAGT 312
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 190 TACAGCGCATGGGAGGAGCAATACGACCGGAGGAGGCTTCTGTGTGTTCGCC 249
QY 313 CTGGAACACCGGAGCTCTTCAGAGAGGTGACAGCGGCTAGGACAGCATC 363
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```


Db 10 TATAGCTGTGTGTGGGCGGCGGTGTGGCAAGCGGCGCTTACATCCAGCTG 69
QY 133 CTCACCGCCGCTTGAGAGCGCTTACAGCGCTTACATCGAGAGCTTCCAGCAAGTTC 192
Db 70 ATCCAGAACCATTTTGTGTGGAGCAATACGACCCCTATAGAGGATTCCTACCGGAGAGC 129
QY 193 TACTCATCCGCGGCGGAGGTCTACAGCTCCGATCTTCGACAGCTCCGCGCAACCCCG 252
Db 130 GTGTGCTATTGATGGAGAGCGTGTGTGGACATCTCGGATTCGCGCGCTGGAGAG 189
QY 253 TTCCCGCGCATCGGCGGCTTCTCATCTCCACAGAGAGAGCTTTTCAT 299
Db 190 TACAGCGCATCGGAGACCATGATGCGACCGGCGGAGGCTTCT 236

RESULT 12
US-08-866A-8

Sequence 8, Application US/08884866A
GENERAL INFORMATION:
APPLICANT: Chien, Shu
APPLICANT: Shyy, John Y-J
TITLE OF INVENTION: GENE THERAPY IN CORONARY ANGIOPLASTY AND
FILE REFERENCE: US081100-1
CURRENT APPLICATION NUMBER: US/08/884,866A
PRIOR FILING DATE: 1997-06-30
PRIOR APPLICATION NUMBER: 60/030,358
NUMBER OF SEQ ID NOS: 25
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 8
LENGTH: 240
TYPE: DNA
ORGANISM: Rattus norvegicus
FEATURE:
NAME/KEY: misc_feature
LOCATION: (49)...(51)
OTHER INFORMATION: nnn = Any nucleic acid triplet, except for UCA,
US-08-884-866A-8
OTHER INFORMATION: UCC, UCU, UGU, AGC, or AGU

Query Match 7.8%; Score 65.6; DB 4; Length 240;
Best Local Similarity 55.1%; Pred. No. 8.5e-05;
Matches 125; Conservative 0; Mismatches 102; Indels 0; Gaps 0;

QY 73 TATGCGATGGTATCCGCGGTGTCTCAAGTGGGCAAGCGGCGCTTACGCTTC 132
Db 10 TATAGCTGTGTGTGGGCGGCGGTGTGGCAANNNGCGCTGACATCCAGCTG 69
QY 133 CTCACCGCGGCTTGAGAGCGCTTACAGCGCTTACATCGAGAGCTTCCAGCAAGTTC 192
Db 70 ATCCAGAACCATTTTGTGTGGAGCAATACGACCCCTATAGAGGATTCCTACCGGAGAGC 129
QY 193 TACTCATCCGCGGCGGAGGTCTACAGCTCCGATCTTCGACAGCTCCGCGCAACCCCG 252
Db 130 GTGTGCTATTGATGGAGAGCGTGTGTGGACATCTCGGATTCGCGCGCTGGAGAG 189
QY 253 TTCCCGCGCATCGGCGGCTTCTCATCTCCACAGAGAGAGCTTTTCAT 299
Db 190 TACAGCGCATCGGAGACCATGATGCGACCGGCGGAGGCTTCT 236

RESULT 13
US-08-884-866A-12
Sequence 12, Application US/08884866A
GENERAL INFORMATION:
APPLICANT: Chien, Shu
APPLICANT: Shyy, John Y-J
TITLE OF INVENTION: GENE THERAPY IN CORONARY ANGIOPLASTY AND
FILE REFERENCE: US081100-1
CURRENT APPLICATION NUMBER: US/08/884,866A
CURRENT FILING DATE: 1997-06-30

PRIOR APPLICATION NUMBER: 60/030,358
PRIOR FILING DATE: 1996-11-08
NUMBER OF SEQ ID NOS: 25
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12
LENGTH: 240
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Variation of SEQ ID NO.:2
US-08-884-866A-12

Query Match 7.7%; Score 65.4; DB 4; Length 240;
Best Local Similarity 55.5%; Pred. No. 9.3e-05;
Matches 126; Conservative 0; Mismatches 101; Indels 0; Gaps 0;

QY 73 TATGCGATGGTATCCGCGGTGTCTCAAGTGGGCAAGCGGCGCTTACGCTTC 132
Db 10 TATAGCTGTGTGTGGGCGGCGGTGTGGCAAGCGGCGCTTACGCTTCAGCTG 69
QY 133 CTCACCGCGGCTTGAGAGCGCTTACAGCGCTTACATCGAGAGCTTCCAGCAAGTTC 192
Db 70 ATCCAGAACCATTTTGTGTGGAGCAATACGACCCCTATAGAGGATTCCTACCGGAGAGC 129
QY 193 TACTCATCCGCGGCGGAGGTCTACAGCTCCGATCTTCGACAGCTCCGCGCAACCCCG 252
Db 130 GTGTGCTATTGATGGAGAGCGTGTGTGGACATCTCGGATTCGCGCGCTGGAGAG 189
QY 253 TTCCCGCGCATCGGCGGCTTCTCATCTCCACAGAGAGAGCTTTTCAT 299
Db 190 TACAGCGCATCGGAGACCATGATGCGACCGGCGGAGGCTTCT 236

RESULT 14
US-08-884-866A-18

Sequence 18, Application US/08884866A
GENERAL INFORMATION:
APPLICANT: Chien, Shu
APPLICANT: Shyy, John Y-J
TITLE OF INVENTION: GENE THERAPY IN CORONARY ANGIOPLASTY AND
FILE REFERENCE: US081100-1
CURRENT APPLICATION NUMBER: US/08/884,866A
PRIOR FILING DATE: 1997-06-30
PRIOR APPLICATION NUMBER: 60/030,358
NUMBER OF SEQ ID NOS: 25
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 18
LENGTH: 240
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Variation of SEQ ID NO.:2
US-08-884-866A-18

Query Match 7.7%; Score 65.4; DB 4; Length 240;
Best Local Similarity 55.5%; Pred. No. 9.3e-05;
Matches 126; Conservative 0; Mismatches 101; Indels 0; Gaps 0;

QY 73 TATGCGATGGTATCCGCGGTGTCTCAAGTGGGCAAGCGGCGCTTACGCTTC 132
Db 10 TATAGCTGTGTGTGGGCGGCGGTGTGGCAAGCGGCGCTTACGCTTCAGCTG 69
QY 133 CTCACCGCGGCTTGAGAGCGCTTACAGCGCTTACATCGAGAGCTTCCAGCAAGTTC 192
Db 70 ATCCAGAACCATTTTGTGTGGAGCAATACGACCCCTATAGAGGATTCCTACCGGAGAGC 129
QY 193 TACTCATCCGCGGCGGAGGTCTACAGCTCCGATCTTCGACAGCTCCGCGCAACCCCG 252
Db 130 GTGTGCTATTGATGGAGAGCGTGTGTGGACATCTCGGATTCGCGCGCTGGAGAG 189
QY 253 TTCCCGCGCATCGGCGGCTTCTCATCTCCACAGAGAGAGCTTTTCAT 299

Db 190 TACAGCGCATGCGGACCAATGCGACCGGGGAGGCTTCT 236

RESULT 15

US-08-866A-19

; Sequence 19, Application US/0884866A

; GENERAL INFORMATION:

; APPLICANT: Chien, Shu

; APPLICANT: Shyy, John Y-J

; TITLE OF INVENTION: GENE THERAPY IN CORONARY ANGIOPLASTY AND

; TITLE OF INVENTION: BYPASS

; FILE REFERENCE: UCSD1100-1

; CURRENT APPLICATION NUMBER: US/08/864,866A

; CURRENT FILING DATE: 1997-06-30

; PRIOR APPLICATION NUMBER: 60/030,358

; PRIOR FILING DATE: 1996-11-08

; NUMBER OF SEQ ID NOS: 25

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 19

; LENGTH: 240

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Variation of SEQ ID NO.:2

; US-08-866A-19

US-08-866A-19

US-08-866A-19

US-08-866A-19

US-08-866A-19

US-08-866A-19

US-08-866A-19

US-08-866A-19

US-08-866A-19

US-08-866A-19

US-08-866A-19

US-08-866A-19

US-08-866A-19

US-08-866A-19

US-08-866A-19

US-08-866A-19

US-08-866A-19

US-08-866A-19

US-08-866A-19

US-08-866A-19

US-08-866A-19

US-08-866A-19

US-08-866A-19

US-08-866A-19

US-08-866A-19

US-08-866A-19

US-08-866A-19

US-08-866A-19

US-08-866A-19

US-08-866A-19

US-08-866A-19

US-08-866A-19

US-08-866A-19

US-08-866A-19

US-08-866A-19

US-08-866A-19

US-08-866A-19

US-08-866A-19

US-08-866A-19

US-08-866A-19

US-08-866A-19

US-08-866A-19

US-08-866A-19

US-08-866A-19

US-08-866A-19

US-08-866A-19

US-08-866A-19

US-08-866A-19

US-08-866A-19

US-08-866A-19

US-08-866A-19

US-08-866A-19

US-08-866A-19

US-08-866A-19

US-08-866A-19

US-08-866A-19

US-08-866A-19

US-08-866A-19

Search completed: December 28, 2002, 05:52:50
Job time: 45.7548 secs

[illegible]

1

Db 168 GCTCACCCGCGTCCACCCGAGGACCCCTCAGCCGCTCTCTGCCCCCTCTCTGCGCCCG 227
 QY 127 GCGCCGCGCTCGCGGCGCCCTCTGCCCCAATGAACCTGGCCGCGATGATCAAGAGATGTC 186
 Db 228 GCGCCGCGCTCGCGGCGCCCTCTGCCCCAATGAACCTGGCCGCGATGATCAAGAGATGTC 287
 QY 187 CCGAGCGACTCGGAGCTGAGTATCCCGGCGCAAGAACTGCTATCGCATGTCATCTCGGC 246
 Db 288 CCGAGCGACTCGGAGCTGAGTATCCCGGCGCAAGAACTGCTATCGCATGTCATCTCGGC 347
 QY 247 TCGTCAAGAGTGGGCAAGACGGCCATCTGTGCGCTTCTCAGCGCGCCGCTTGAGAGAC 306
 Db 348 TCGTCAAGAGTGGGCAAGACGGCCATCTGTGCGCTTCTCAGCGCGCCGCTTGAGAGAC 407
 QY 307 GCTACAGCGCTACCATGAGAGACTTCTCCACCGAAGTCTCTACTCCATCCGCGCGAGTTC 366
 Db 408 GCTACAGCGCTACCATGAGAGACTTCTCCACCGAAGTCTCTACTCCATCCGCGCGAGTTC 467
 QY 367 TACAGCTCGACATCTCTGACACAGTCCGCGCAACACCGCTTCCCGCCATGCGCGCGCTC 426
 Db 468 TACAGCTCGACATCTCTGACACAGTCCGCGCAACACCGCTTCCCGCCATGCGCGCGCTC 527
 QY 427 TCCATCTCTACAGAGAGAGTCTTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 486
 Db 528 TCCATCTCTACAGAGAGAGTCTTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 587
 QY 487 GAGGAGTTCACAGCGGCTCAGCAGAGATCTCGACACCAAGTCTTGGCTCAGAGACAA 546
 Db 588 GAGGAGTTCACAGCGGCTCAGCAGAGATCTCGACACCAAGTCTTGGCTCAGAGACAA 647
 QY 547 ACCAAGAGAGAGTGGAGCTGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 606
 Db 648 ACCAAGAGAGAGTGGAGCTGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 707
 QY 607 TACCGCGAGTGGAGCAGCAGCGGAGATCGAGCGAGTGGGCGGAGACCGCCGAGCGCTGC 666
 Db 708 TACCGCGAGTGGAGCAGCAGCGGAGATCGAGCGAGTGGGCGGAGACCGCCGAGCGCTGC 767
 QY 667 GCTACTTCGAGATCTCGGCGCAAGAAAGACAGCAGCTGGACAGATGTTCCGCGCGCTC 726
 Db 768 GCTACTTCGAGATCTCGGCGCAAGAAAGACAGCAGCTGGACAGATGTTCCGCGCGCTC 827
 QY 727 TTTGCGCATGGCCAGAGTGGCCGCGAGATGAGCCAGACCTGACCGCAAGGTCTCGGTG 786
 Db 828 TTTGCGCATGGCCAGAGTGGCCGCGAGATGAGCCAGACCTGACCGCAAGGTCTCGGTG 887
 QY 787 CAGTACTGGAGCTGCTGCTGACAGAAAGGCGCTGCGGAACAAAGACTGCTGCGGCGCGC 846
 Db 888 CAGTACTGGAGCTGCTGCTGACAGAAAGGCGCTGCGGAACAAAGACTGCTGCGGCGCGC 947
 QY 847 AGC 906
 Db 948 AGC 1007
 QY 907 CCGCGCGAGCTGACAGCGCACTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 966
 Db 1008 CCGCGCGAGCTGACAGCGCACTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1067
 QY 967 GCCAAG 1026
 Db 1068 GCCAAG 1127
 QY 1027 TAAG 1086
 Db 1128 TAAG 1187
 QY 1087 CCGCGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1146
 Db 1188 CCGCGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1247
 QY 1147 AACCGAG 1206
 Db 1248 AACCGAG 1307

QY 1207 TTCCCGACCCCGGCTCCCGCATTTAGAGCCCGCACCCCGCAATACCTTTGGAGCGAGGCGC 1266
 Db 1308 TTCCCGACCCCGGCTCCCGCATTTAGAGCCCGCACCCCGCAATACCTTTGGAGCGAGGCGC 1366
 QY 1267 CAGCCGAGGCTGATTTATCTTCTCAAAAGACCTTAAGATGAGCGCGGCTGGGGAGGGA 1326
 Db 1367 CAGCCGAGGCTGATTTATCTTCTCAAAAGACCTTAAGATGAGCGCGGCTGGGGAGGGA 1426
 QY 1327 TGTGAATTTATCAGCGCTTCTGCTAGAGCTTCAAGAAACCGTCATGCCGCTTGAGGCTCAG 1386
 Db 1427 TGTGAATTTATCAGCGCTTCTGCTAGAGCTTCAAGAAACCGTCATGCCGCTTGAGGCTCAG 1486
 QY 1387 GACCCAGGCGGCGATTTATCTTCTGCTAGATTCGCGGTTGCTGTGACAGCGCGTGAAGCTTC 1446
 Db 1487 GACCCAGGCGGCGATTTATCTTCTGCTAGATTCGCGGTTGCTGTGACAGCGCGTGAAGCTTC 1546
 QY 1447 TGCCCTCCCGAAGTAAAGCGGCGGCGGCTGGGTCAATCATTAACCAAGTACTGTTTAC 1506
 Db 1547 TGCCCTCCCGAAGTAAAGCGGCGGCGGCTGGGTCAATCATTAACCAAGTACTGTTTAC 1606
 QY 1507 ATGTAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1566
 Db 1607 ATGTAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1666
 QY 1567 GTGTCAATGAGACAGCAAGCAAAACCTTACCAGGTGTTTAACTGTGTGTAGGT 1626
 Db 1667 GTGTCAATGAGACAGCAAGCAAAACCTTACCAGGTGTTTAACTGTGTGTAGGT 1726
 QY 1627 CTTTAAAGTATTTGCTTATTTGTTTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 1686
 Db 1727 CTTTAAAGTATTTGCTTATTTGTTTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 1786
 QY 1687 AAAAAAAAAA 1698
 Db 1787 AAAAAAAAAA 1798

RESULT 2
 US-09-53-374A-3
 : Sequence 3, Application US/09053374A
 : Patent No. 6462177
 :
 : GENERAL INFORMATION:
 : APPLICANT: YEN, KWANG-MU
 : TITLE OF INVENTION: MAMMALIAN BLOOD LOSS-INDUCED GENE, KD312
 : NUMBER OF SEQUENCES: 9
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: AMGEN INC.
 : STREET: ONE AMGEN CENTER DRIVE
 : CITY: THOUSAND OAKS
 : STATE: CA
 : COUNTRY: US
 : ZIP: 91320
 :
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: Patentin Release #1.0, Version #1.30
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/09/053,374A
 : FILING DATE:
 : CLASSIFICATION: 435
 : ATTORNEY/AGENT INFORMATION:
 : NAME: COOK, ROBERT R.
 : REGISTRATION NUMBER: 31,602
 : REFERENCE/DOCKET NUMBER: A-514
 : INFORMATION FOR SEQ ID NO: 3:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 3986 base pairs
 : TYPE: nucleic acid
 : STRANDEDNESS: single
 : TOPOLOGY: linear
 : MOLECULE TYPE: DNA (genomic)


```

; Sequence 9, Application US/08884866A
; GENERAL INFORMATION:
; APPLICANT: Chien, Shu
; APPLICANT: Shyy, John Y-J
; TITLE OF INVENTION: GENE THERAPY IN CORONARY ANGIOPLASTY AND
; FILE REFERENCE: BYPASS
; FILE REFERENCE: UCS01100-1
; CURRENT APPLICATION NUMBER: US/08/884,866A
; CURRENT FILING DATE: 1997-06-30
; PRIOR APPLICATION NUMBER: 60/030,358
; PRIOR FILING DATE: 1996-11-08
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 480
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (49)...(51)
; OTHER INFORMATION: nnn = Any nucleic acid triplet, except for UCA,
; OTHER INFORMATION: UCC, UCG, UCU, AGC, or AGU
US-08-884-866A-9

```

```

RESULT 5
US-08-884-866A-10
; Sequence 10, Application US/08884866A
; GENERAL INFORMATION:
; APPLICANT: Chien, Shu
; APPLICANT: Shyy, John Y-J
; TITLE OF INVENTION: GENE THERAPY IN CORONARY ANGIOPLASTY AND
; FILE REFERENCE: BYPASS
; FILE REFERENCE: UCS01100-1
; CURRENT APPLICATION NUMBER: US/08/884,866A
; CURRENT FILING DATE: 1997-06-30
; PRIOR APPLICATION NUMBER: 60/030,358
; PRIOR FILING DATE: 1996-11-08
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 450
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (49)...(51)
; OTHER INFORMATION: nnn = Any nucleic acid triplet, except for UCA,
; OTHER INFORMATION: UCC, UCG, UCU, AGC, or AGU
US-08-884-866A-10

```

```

Query Match 5.0%; Score 89.6; DB 4; Length 450;
Best Local Similarity 56.4%; Pred. No. 1.1e-07;
Matches 164; Conservative 0; Mismatches 127; Indels 0; Gaps 0;

```

```

QY 226 TATCGATGTCATCTCGGCTGTCCTCAAGTGGGCAAGACGGCCATCGTGGCGCTTC 285
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 10 TATAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 69
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 286 CTGACCGCGCGCTTCGAGAGCGCTTACAGCGCTTACAGAGACTTCCACCGCAAGTTC 345
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 70 ATCCAGAACCATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 129
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 346 TACTCATCGCGCGCGCGCGCTTACAGAGCTTACAGAGCTTCCAGAGCGCAAGCGG 405
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 130 GTGTGTATTTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 189
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 406 TTCCCGCGCGCGCGCGCGCTTCCATCTTCACAGAGAGCTTTTCATCTCTGTGTTCAGT 465
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 190 TACAGCGCGCGCGCGCGCGAGTCAATGCGCACCGGGGAGGCGCTTCCGTTGTGTGCC 249
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 466 CTGACAGACCGCGAGCTCTTCGAGAGGTGCAAGCGGTCTGAGAGAGATC 516
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 250 ATCAACAACACCAAGTCTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 300
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 6
US-08-884-866A-9

```

```

; Sequence 9, Application US/08884866A
; GENERAL INFORMATION:
; APPLICANT: Chien, Shu
; APPLICANT: Shyy, John Y-J
; TITLE OF INVENTION: GENE THERAPY IN CORONARY ANGIOPLASTY AND
; FILE REFERENCE: BYPASS
; FILE REFERENCE: UCS01100-1
; CURRENT APPLICATION NUMBER: US/08/884,866A
; CURRENT FILING DATE: 1997-06-30
; PRIOR APPLICATION NUMBER: 60/030,358
; PRIOR FILING DATE: 1996-11-08
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 480
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (49)...(51)
; OTHER INFORMATION: nnn = Any nucleic acid triplet, except for UCA,
; OTHER INFORMATION: UCC, UCG, UCU, AGC, or AGU
US-08-884-866A-9

```

```

Query Match 5.0%; Score 89.6; DB 4; Length 480;
Best Local Similarity 56.4%; Pred. No. 1.1e-07;
Matches 164; Conservative 0; Mismatches 127; Indels 0; Gaps 0;

```

```

QY 226 TATCGATGTCATCTCGGCTGTCCTCAAGTGGGCAAGACGGCCATCGTGGCGCTTC 285
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 10 TATAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 69
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 286 CTGACCGCGCGCTTCGAGAGCGCTTACAGCGCTTACAGAGACTTCCACCGCAAGTTC 345
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 70 ATCCAGAACCATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 129
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 346 TACTCATCGCGCGCGCGCGCTTACAGAGCTTACAGAGCTTCCAGAGCGCAAGCGG 405
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 130 GTGTGTATTTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 189
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 406 TTCCCGCGCGCGCGCGCGCTTCCATCTTCACAGAGAGCTTTTCATCTCTGTGTTCAGT 465
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 190 TACAGCGCGCGCGCGCGAGTCAATGCGCACCGGGGAGGCGCTTCCGTTGTGTGCC 249
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 466 CTGACAGACCGCGAGCTCTTCGAGAGGTGCAAGCGGTCTGAGAGAGATC 516
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 250 ATCAACAACACCAAGTCTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 300
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```



```

RESULT 10
US-08-545-196B-10
; Sequence 10, Application US/08545196B
; Patent No. 6080577
; GENERAL INFORMATION:
; APPLICANT: MELKI, JUDITH
; APPLICANT: MUNNICH, ARNOLD
; TITLE OF INVENTION: SURVIVAL MOTOR NEURON (SMN) GENE: A GENE
; TITLE OF INVENTION: FOR SPINAL MUSCULAR ATROPHY
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESS: BIRCH, STEWART, KOLASCH AND BIRCH, LLP
; STREET: PO BOX 747
; CITY: FALLS CHURCH
; STATE: VA
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/545,196B
; FILING DATE: 19-OCT-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: FARACI, C. J.
; REGISTRATION NUMBER: 32,350
; REFERENCE/DOCKET NUMBER: 2121-110P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1582 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-545-196B-10

Query Match          4.7%   Score 85.2: DB 3: Length 1582:
Best Local Similarity 74.0%; Pred No.8.8e-07;
Matches 108; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 1602 TGTTTACTGTCGTGTGTCGAGCTTAAAGTAATTCCTTAATTGGCTTTTAAATAT 1661
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1437 TTTCATCTCTCAATAAGTTAAAGATATTAATAAATATTTATTTTTTAAAAA 1496

QY 1662 ACAATAAATATTTAAAGAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1721
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1497 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1556

QY 1722 AAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1747
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1557 AAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1582

RESULT 11
US-08-545-196B-12
; Sequence 12, Application US/08545196B
; Patent No. 6080577
; GENERAL INFORMATION:
; APPLICANT: MELKI, JUDITH
; APPLICANT: MUNNICH, ARNOLD
; TITLE OF INVENTION: SURVIVAL MOTOR NEURON (SMN) GENE: A GENE
; TITLE OF INVENTION: FOR SPINAL MUSCULAR ATROPHY
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESS: BIRCH, STEWART, KOLASCH AND BIRCH, LLP
; STREET: PO BOX 747
; CITY: FALLS CHURCH
; STATE: VA
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/545,196B
; FILING DATE: 19-OCT-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: FARACI, C. J.
; REGISTRATION NUMBER: 32,350
; REFERENCE/DOCKET NUMBER: 2121-110P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1582 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA

```

1 STREET: PO BOX 747
2 CITY: FALLS CHURCH
3 STATE: VA
4
5 COUNTRY: USA
6 ZIP: 22040-0747
7
8 COMPUTER READABLE FORM:
9 MEDIUM TYPE: Floppy disk
10 COMPUTER: IBM PC compatible
11 OPERATING SYSTEM: PC-DOS/MS-DOS
12
13 SOFTWARE: PatentIn Release #1.0, Version #1.30
14
15 CURRENT APPLICATION DATA:
16 APPLICATION NUMBER: US/08/545,196B
17 FILING DATE: 19-OCT-1995
18 CLASSIFICATION: 435
19
20 ATTORNEY/AGENT INFORMATION:
21
22 NAME: PARACI, C. J.
23 REGISTRATION NUMBER: 32,350
24
25 REFERENCE/DOCKET NUMBER: 2121-110P
26
27 TELECOMMUNICATION INFORMATION:
28 TELEPHONE: (703) 205-8000
29 TELEFAX: (703) 205-8050
30
31 INFORMATION FOR SEQ ID NO: 12:
32
33 SEQUENCE CHARACTERISTICS:
34
35 LENGTH: 1582 base pairs
36 TYPE: nucleic acid
37 STRANDEDNESS: double
38
39 TOPOLOGY: linear
40
41 MOLECULE TYPE: CDNA
42
43 US-08-545-196B-12

```

Query Match 4.7%; Score 85.2; DB 3; Length 1582;
Best Local Similarity 74.0%; Pred. No. 8.8e-07;
Matches 108; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

OY 1602 TGTTCATACGTGTCGTGTCGAGGCTTTAAAGTTATGCTTTATTTGGTTTAAATAT 1661
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1437 TTTCTATCTCTCTATATGTTTAAAGTATATATATAATATTTATTTTAAAAA 1496
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |

OY 1662 ACAATTAATATATTTAAATGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1721
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1497 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1556
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |

OY 1722 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1747
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1557 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1582
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 12
US-09-370-253-1
; Sequence 1, Application US/09370253
; Patent No. 6165792
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen M.
; APPLICANT: Sakai, Hajime
; APPLICANT: Thorpe, Catherine J.
; TITLE OF INVENTION: Amino Acid Transporters
; FILE REFERENCE: BB-1200
; CURRENT APPLICATION NUMBER: US/09/370,253
; CURRENT FILING DATE: 1999-08-09
; EARLIER APPLICATION NUMBER: 60/097,222
; EARLIER FILING DATE: August 20, 1998
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 1
; LENGTH: 1882
; TYPE: DNA
; ORGANISM: Hordeum vulgare
US-09-370-253-1

Query Match 4.5%; Score 81.8; DB 4; Length 1882;
Best Local Similarity 81.2%; Pred. No. 3.7e-06;
Matches 95; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

```



```

Db 2442 TGTAAATCTATTTATGTTTATTTTGTAAATTAAGATTCTTTTAACCACTGGCA 2501
QY 1662 ACAATTAATATATTTTAAATGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1721
Db 2502 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2561
QY 1722 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCGGC 1754
Db 2562 AAAAAAAAAAAAAAAAAAAAAAAAAAGGGGGGC 2594

```

Search completed: December 28, 2002, 05:53:08
 Job time : 96.2452 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: December 28, 2002, 01:13:57 ; Search time 76 Seconds

(without alignments)
1133.897 Million cell updates/sec

Title: US-09-709-103-2
Perfect score: 1449

Sequence: 1 MKLAAMIKKMCPSDSELSIP.....IRKASAGQADKRCVCYS 281

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 441362 segs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL-frame+p2n -DEV-rlp
-O/cg92.1/USFTO.spool/US09709103/runat.20122002.162406.833/app-query.fasta.1.455
-DB-Issued_Patents_NA -QEMT-fastap -SUFFIX-p2n.rn1 -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS-bits -START=1 -END=1 -MATRIX-biosum62 -TRANS-human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pcr -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE-LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USR=US09709103 @CGN.1.1.31 @runat.20122002.162406.833 -NCTU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV-TIMEOUT=120
-WARN_TIMEOUT=30 -TREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Issued_Patents_NA:*

1: /cg92.6/ptodata/1/lna/5A.COMB.seq:*
2: /cg92.6/ptodata/1/lna/5B.COMB.seq:*
3: /cg92.6/ptodata/1/lna/6A.COMB.seq:*
4: /cg92.6/ptodata/1/lna/6B.COMB.seq:*
5: /cg92.6/ptodata/1/lna/PCRTUS.COMB.seq:*
6: /cg92.6/ptodata/1/lna/Backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1449	100.0	1841	4	US-09-053-374A-1
2	1399.5	96.6	1689	4	US-09-053-374A-4
3	1387	95.7	3986	4	US-09-053-374A-3
4	1343	92.7	3079	4	US-09-053-374A-6
5	300	20.7	615	1	US-08-247-946A-5
6	300	20.7	615	5	PCT-US95-06420-5
7	279	19.3	5775	1	US-08-306-691B-15
8	279	19.3	5775	5	PCT-US93-06251-29
9	277.5	19.2	570	4	US-08-884-866A-2
10	277.5	19.2	570	4	US-08-884-866A-11
11	276	19.0	607	2	US-08-429-96A-85
12	275.5	19.0	480	4	US-08-884-866A-9

13	271.5	18.7	4480	4	US-09-167-322-12
14	266.5	18.4	450	4	US-08-884-866A-10
15	266.5	18.4	574	2	US-08-429-96A-83
16	265.5	18.3	2436	1	US-08-306-691B-16
17	237	16.4	2309	3	US-09-078-317-3
18	237	16.4	2309	3	US-09-454-818-3
19	234.5	16.2	600	3	US-09-078-317-1
20	234.5	16.2	600	4	US-09-454-818-1
21	219.5	15.1	563	4	US-09-454-818-1
22	204	14.1	3497	4	US-09-385-982-426
23	186	12.8	6453	1	US-08-503-505A-2
24	186	12.8	6453	3	US-08-306-691B-14
25	186	12.8	6453	3	US-09-209-668-10
26	185.5	12.8	2964	2	US-09-356-952-8
27	185.5	12.8	2964	3	US-08-846-790A-2
28	180	12.4	702	3	US-08-935-333-2
29	180	12.4	702	3	US-08-842-975-2
30	180	12.4	702	3	US-09-213-397-2
31	177.5	12.2	1166	5	US-09-416-489-2
32	175.5	12.1	603	4	PCT-US96-12129B-1
33	175.5	12.1	932	4	US-09-325-932A-29
34	175	12.1	1175	4	US-09-325-932A-28
35	173	11.9	1175	4	US-09-387-341-215
36	172	11.9	1175	2	US-08-773-423-6
37	172	11.9	897	2	US-09-006-535-2
38	172	11.9	1443	1	US-08-076-089-1
39	172	11.9	1443	2	US-08-707-200-1
40	172	11.9	1443	1	US-08-996-565-1
41	172	11.9	1443	5	US-08-996-565-1
42	167	11.5	1525	2	PCT-US93-05643-1
43	166.5	11.5	897	2	US-09-006-535-1
44	166	11.5	5197	4	US-09-006-535-7
45	165	11.4	890	4	US-09-293-170-6
					US-09-075-454-14
					US-08-884-866A-8

ALIGNMENTS

RESULT 1
US-09-053-374A-1
; Sequence 1, Application US/09053374A
; Patent No. 6462177
; GENERAL INFORMATION:
; APPLICANT: YEN, KWANG-MU
; TITLE OF INVENTION: MAMMALIAN BLOOD LOSS-INDUCED GENE, KD312
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: AMGEN INC.
; STREET: ONE AMGEN CENTER DRIVE
; CITY: THOUSAND OAKS
; STATE: CA
; COUNTRY: US
; ZIP: 91320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/053,374A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: COOK, ROBERT R. 31,602
; REGISTRATION NUMBER: A-514
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1841 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:

NAME/KEY: CDS
LOCATION: 255..1097
US-09-053-374A-1

Alignment Scores:

Pred. No.: 5,33e-166 Length: 1841
Score: 1449.00 Matches: 281
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-09-709-103-2 (1-281) x US-09-053-374A-1 (1-1841)

QY 1 MetLysLeuAlaAlaMetLysLysMetCysProSerAspSerGluLeuSerIlePro 20
DB 255 ATGAAACTGGCGCGATGATCAAGAAAGATGCGCCGAGCTCGAGTGGATCCCG 314
QY 21 AlaLysAsnCysTyrArgMetValIleLeuGlySerSerValGlyLysThrAlaIle 40
DB 315 GCCAAGACTGCTATGCGATGCGATCCCGGCTGCTCCAGAGTGGGCAAGCGCGCATC 374
QY 41 ValSerArgPheLeuThrGlyArgPheGluAspAlaTyrThrProThrIleGluAspPhe 60
DB 375 GTGTCCGCTCTCTCAACCGCGCTTCGAGGAGCGCTTACACGCTTACATCGAGGACTTC 434
QY 61 HisArgLysPheTyrSerIleArgGlyGluValTyrGlnLeuAspIleLeuAspThrSer 80
DB 435 CACCGCAAGTGTCTACTCTCCCGCGGAGGTCTACCACTGCACTCTCCGACAGCTCC 494
QY 81 GlyAsnHisProPheProAlaMetArgArgLysSerIleLeuThrGlyAspValPheIle 100
DB 495 GGCACACACCGCTTCCCGCGCATGCGCGCTCTCCATCTCCACAGAGAGCGTTTCATC 554
QY 101 LeuValPheSerLeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGln 120
DB 555 CTGGTGTGCTGTGCAACCGCGACTCTTCGACAGAGTGGAGCGGCTCGACGCGAG 614
QY 121 IleLeuAspThrLysSerCysLeuLysAsnLysThrLysGluAsnValAspValProLeu 140
DB 615 ATCTCTGCACCAAGTCTTCTCCCAAGAACCAAGAGAGAGTGGACGTGCCCTG 674
QY 141 ValIleCysGlyAsnLysGlyAspArgAspPheTyrArgGluValAspGlnArgGluIle 160
DB 675 GTCATCTGCGGCAACAGGAGTGGACCGGACTTCTACCGGAGGTGGAGCAGCGCGAGATC 734
QY 161 GluGlnLeuValGlyAspAspProGlnArgCysAlaTyrPheGluIleSerAlaLysLys 180
DB 735 GAGCAGCTGTGGGCGAGCGACCGCGCTGCTACTTCGAGATCTGGCGCAAGAG 794
QY 181 AsnSerSerLeuAspGlnMetPheArgAlaLeuPheAlaMetAlaLysLeuProSerGlu 200
DB 795 AACGAGCGCTGGACGAGATGTTCCGCGCTTCCGCAATGGCCAAAGTGGCCACGAG 854
QY 201 MetSerProAspLeuHisArgLysValSerValGlnTyrCysAspValLeuHisLysLys 220
DB 855 ATGAGCCCAACCTGCACCGCCCAAGGTCTCGGTGCAAGTGCAGCGTGCACAGAGAG 914
QY 221 AlaLeuArgAsnLysLysLeuLeuArgAlaGlySerGlyGlyGlyGlyGlyAspProGly 240
DB 915 GCGCTGGGAAACAGAGAGCTGCTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 974
QY 241 AspAlaPheGlyIleValAlaProPheAlaArgArgProSerValHisSerAspLeuMet 260
DB 975 GAGCGCTTGGCATCGTGGACCGCTTCCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1034
QY 261 TyrIleArgGlyLysLysAlaSerAlaGlySerGlnAlaLysAspLysGluArgCysValIle 280
DB 1035 TACATCGCGGAGAGCGCGAGCGCGGAGCGCGCAAGGACAGAGAGCGCTGCTATC 1094
QY 281 Ser 281
DB 1095 AGC 1097

RESULT 2

US-09-053-374A-4

Sequence 4, Application US/09053374A

Patent No. 6462177

GENERAL INFORMATION:

APPLICANT: YEN, KWANG-MU

TITLE OF INVENTION: MAMMALIAN BLOOD LOSS-INDUCED GENE, KD312

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:

ADDRESSEE: AMGEN INC.

STREET: ONE AMGEN CENTER DRIVE

CITY: THOUSAND OAKS

STATE: CA

COUNTRY: US

ZIP: 91320

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/053,374A

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: COOK, ROBERT R.

REGISTRATION NUMBER: 31,602

REFERENCE/DOCKET NUMBER: A-514

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 1689 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 132..971

US-09-053-374A-4

Alignment Scores:

Pred. No.: 4.74e-160 Length: 1689
Score: 1399.50 Matches: 273
Percent Similarity: 98.22% Conservative: 3
Best Local Similarity: 97.15% Mismatches: 4
Query Match: 96.58% Indels: 1
DB: Gaps: 1

US-09-709-103-2 (1-281) x US-09-053-374A-4 (1-1689)

QY 1 MetLysLeuAlaAlaMetLysLysMetCysProSerAspSerGluLeuSerIlePro 20
DB 132 ATGAAACTGGCGCGATGATCAAGAAAGATGCGCCGAGCTCGAGTGGATCCCG 191
QY 21 AlaLysAsnCysTyrArgMetValIleLeuGlySerSerValGlyLysThrAlaIle 40
DB 192 GCCAAGACTGCTATGCGATGCGATCCCGGCTATCCCAAGTGGGCAAGCGCGCATC 251
QY 41 ValSerArgPheLeuThrGlyArgPheGluAspAlaTyrThrProThrIleGluAspPhe 60
DB 252 GTGTCCGCTCTCTCAACCGCGCTTCGAGGAGCGCTTACACCGCTTACATCGAGGACTTC 311
QY 61 HisArgLysPheTyrSerIleArgGlyGluValTyrGlnLeuAspIleLeuAspThrSer 80
DB 312 CACCGCAAGTGTCTACTCTCCCGCGGAGGTCTACCACTGCACTCTCCGACAGCTCT 371
QY 81 GlyAsnHisProPheProAlaMetArgArgLysSerIleLeuThrGlyAspValPheIle 100
DB 372 GAGCATCATCTGCTTCCCGCATGCGCGCTCTCTTACTCTACAGGAGAGCTTTTCATT 431
QY 101 LeuValPheSerLeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGln 120

Db 432 CTGGTGTTCAGCTTAGACACCGGACGCTTCGAGAGAGTGCACAAAGGCTCAACACAGCAG 491
 QY 121 ILEUASPHTLysSerCysLeuLysAsnLysThrLysGluAsnValAspValProLeu 140
 Db 492 ATCTAGACACCAAGTCTCTCTCAGAAACCAAGAGATGTGGAGCTCCGCTG 551
 QY 141 ValIleCysGlyAsnLysGlyAspArgAspPheTyrTrpGluValAspGlnArgGluIle 160
 Db 552 GTCAATTCGGGTAAACAAGGGGACCGGACCTTCAACCCGCAAGAGGAGACGCGGAGATTT 611
 QY 161 GluGlnLeuValGlyAspAspProGlnArgCysAlaTyrPheGluIleSerAlaLysLys 180
 Db 612 GAGACAGCGGTGGGCGATGAGACCCGACGCTGTCTCTCAGATCTCGCCCAAGAG 671
 QY 181 AsnSerSerLeuAspGlnMetPheArgAlaLeuPheAlaMetAlaLysLeuProSerGlu 200
 Db 672 AATGACGACCTGGACACGAGATGTTCCGTGCGCTTTGCTCCATGGCCCAAGCTGCTAGCGAG 731
 QY 201 MetSerProAspLeuHisArgLysValSerValGlnTyrCysAspValIleuHisLysLys 220
 Db 732 ATGAGCCCTGACTTGCACCCGCAAGGTCTGTGTGCACTGTGACGTCGTCACAAAG 791
 QY 221 AlaLeuArgAsnLysLysLeuLeuArgAlaGlySerGlyGlyGlyGlyAspProGly 240
 Db 792 GCTCTGAGGAACAAGAGCTTCTGCGTGGCGGAGC---GGAGGTGGGGGCGACACGGA 848
 QY 241 AspAlaPheGlyIleValAlaProPheAlaArgArgProSerValHisSerAspLeuMet 260
 Db 849 GATGCTTTGGCATCTGGGCGCCCTTGTCTGCGACGCTGCTGCTATGCGACCTCATG 908
 QY 261 TyrIleArgLysLysAlaSerAlaGlySerGlnAlaLysAspLysGluArgCysValIle 280
 Db 909 TACATTGCTGAGAAACCACTGTCAGCAGCAGGCTTAAGACAAAGGAGCGCTGTGTATC 968
 QY 281 Ser 281
 Db 969 AGT 971

RESULT 3

US-09-053-374A-3
 Sequence 3, Application US/09053374A
 Patent No. 6462177
 GENERAL INFORMATION:
 APPLICANT: YEN, KWANG-MU
 TITLE OF INVENTION: MAMMALIAN BLOOD LOSS-INDUCED GENE, KD312
 NUMBER OF SEQUENCES: 9
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: AMGEN INC.
 STREET: ONE AMGEN CENTER DRIVE
 CITY: THOUSAND OAKS
 STATE: CA
 COUNTRY: US
 ZIP: 91320
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/053,374A
 FILING DATE:
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: COOK, ROBERT R.
 REGISTRATION NUMBER: 31,602
 REFERENCE/DOCKET NUMBER: A-514
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3986 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)

US-09-053-374A-3
 Alignment Scores:
 Pred. No.: 5 53e-158
 Score: 1387.00
 Percent Similarity: 79.83%
 Best Local Similarity: 79.83%
 Query Match: 95.72%
 DB: 4
 Gaps: 1
 US-09-709-103-2 (1-281) x US-09-053-374A-3 (1-3986)
 QY 1 MetLysLeuAlaAlaMetIleLysLysMetCysProSerAspSerGluLeuSerIlePro 20
 Db 774 ATGAACCTGGCCCGCATGATCAAGAAAGATGTGCCAGGAGCTCGAGCTGATTCGG 833
 QY 21 AlaLysAsnCysTyrArgMetValIleLeuGlySerSerLysValGlyLysThrAlaIle 40
 Db 834 GCCAAGACTGCTATGTCATGTGTCATCTCGGCTGCTCCAAAGTGGGCAAGAGCCATC 893
 QY 41 ValSerArgPheLeuThrGlyArgPheGluAspAlaTyrThrProThrIleGluAspPhe 60
 Db 894 GTCTCGCGCTTCTCTACCGCGCCGCTTCGAGAGCGCTACACGCTTACACCTGAGGACTTC 953
 QY 61 HisArgLysPheTyrSerIleArgGlyGluValTyrGlnLeuAspIleLeuAspThrSer 80
 Db 954 CACCGCAAGTTCTACTCCATCCGCGGAGGTCCTACACACTGACATCTCGACACGCTCC 1013
 QY 81 GlyAsnHisProPheProAlaMetArgArgLeuSerIleLeuThr----- 95
 Db 1014 GGCACACACCCCTTCCCGCATGGGCGGCGCTCTCATCTCCACAGTGGAGCGGGGGCC 1073
 QY 95 ----- 95
 Db 1074 GGGCAGTGGGGAGGAGGAGGGGCGGGAACCTCCGCGCAAGGCGCCCGGAGCGCGCT 1133
 QY 95 ----- 95
 Db 1134 CCGGCTGTGCGCGCGAGTAGTGCCTTCCGCTTAGAGAGCTAAGCGCGCCCGCGCG 1193
 QY 95 ----- 95
 Db 1194 GCCTCAAGTACAGCCGACTTGTCCCTGGGGGCGCACCTCACCCTTCTCTTCTGCT 1253
 QY 96 -----GlyAspValPheIleuValPheSerLeuAspAsnArgAspSe 110
 Db 1254 CTCTGCCCCCTCTAGAGAGAGCTTTTCACTCTGCTGCTTCACTGCAACAACCGGACTC 1313
 QY 110 PheGluGluValGlnArgLeuArgGlnIleLeuAspThrLysSerCysLeuLysAs 130
 Db 1314 CTTGAGGAGAGTGCAGCGGCTCAGCAGACAGATCCTGCACACCAAGTCTTGCCTCAAGAA 1373
 QY 130 nLysThrLysGluAsnValAspValProLeuValIleCysGlyAsnLysGlyAspArgAs 150
 Db 1374 CAACAACCAAGGAGAGAGGTGAGCTGCCCTGGTCACTGCGGCAACAAGGTGAGCCGGA 1433
 QY 150 PheTyrArgLysValAspGlnArgGluIleGluGlnLeuValGlyAspAspProGlnArg 170
 Db 1434 CTTTACCGCGAGGTGAGCAAGCGGAGATGAGAGAGCTGTGGGCGACGACCCCAAGC 1493
 QY 170 GlysAlaTyrPheGluIleSerAlaLysLysAsnSerSerLeuAspGlnMetPheArgAl 190
 Db 1494 CTGCGCTTACTTTCGAGATCTCGGCCCAAGAGACGAGCTTGAACCGAGTTCGCGGC 1553
 QY 190 AlaPheAlaMetAlaLysLeuProSerGlnMetSerProAspLeuHisArgLysValSe 210
 Db 1554 GCTCTTGGCATGGCCAGAGCTGCCCCAGGAGATGAGCCCAACCTGACCCCAAGGCTC 1613
 QY 210 TValGlnTyrCysAspValIleuHisLysLysAlaLeuArgAsnLysLysLeuLeuArgAl 230
 Db 1614 GGTGAGTACTGCGACGTCGTGCAACAAGAGCGCTGCGGAGCAACAAGAGCTGCGCGGC 1673
 QY 230 aGlySerGlyLysGlyGlyAspProGlyAspAlaPheGlyIleValAlaProPheAl 250

Db	1674	CGGAGAGGGGGGGGGGGGGCGACCGGCGAGCGCTTGGCATGCTGGACACCTTCGC	1733
Qy	250	aatgagproservalhisserapleumetyrlleahrglulysalaserAlaglyse	270
Db	1734	GGCGGGGGCCGAGCTTACAGCGACGACCTCAATGTCATTCGCGAGAAAGCCAGCGCGGAG	1793
Qy	270	rglAlalalsaplysgluarqCysvalIleSer	281
Db	1794	CCAGGCCAAGACAGAGAGGCTGCGTCATCAGC	1827
RESULT 4			
US-09-053-374A-6			
US-09-053-374A-6			
Sequence 6, Application US/09053374A			
Patent No. 6462177			
GENERAL INFORMATION:			
APPLICANT: YEN, KWANG-MU			
TITLE OF INVENTION: MAMMALIAN BLOOD LOSS-INDUCED GENE, KD312			
NUMBER OF SEQUENCES: 9			
CORRESPONDENCE ADDRESS:			
ADDRESSEE: AMGEN INC.			
STREET: ONE AMGEN CENTER DRIVE			
CITY: THOUSAND OAKS			
STATE: CA			
COUNTRY: US			
ZIP: 91320			
COMPUTER READABLE FORM:			
MEDIUM TYPE: Floppy disk			
COMPUTER: IBM PC compatible			
OPERATING SYSTEM: PC-DOS/MS-DOS			
SOFTWARE: Patentn Release #1.0, Version #1.30			
CURRENT APPLICATION DATA:			
APPLICATION NUMBER: US/09/053,374A			
FILING DATE:			
CLASSIFICATION: 435			
ATTORNEY/AGENT INFORMATION:			
NAME: COOK, ROBERT R.			
REGISTRATION NUMBER: 31,602			
REFERENCE/DOCKET NUMBER: A-514			
INFORMATION FOR SEQ ID NO: 6:			
SEQUENCE CHARACTERISTICS:			
LENGTH: 3079 base pairs			
TYPE: nucleic acid			
STRANDEDNESS: single			
TOPOLOGY: linear			
MOLECULE TYPE: DNA (genomic)			
US-09-053-374A-6			
Alignment Scores:			
Pred. No.: 8,2e-153			
Score: 1343.00			
Percent Similarity: 84.10%			
Best Local Similarity: 83.18%			
Query Match: 92.68%			
DB: 4			
Gaps: 2			
US-09-709-103-2 (1-281) x US-09-053-374A-6 (1-3079)			
Qy	1	MettlySleuAlaAlaMeIllelyslslyMetySproseraspsergluSersIlepro	20
Db	303	ATGAAACAGCGCGCATGATCAAGAAAGTGTGCCCAAGCAGCTTCGAGCTATCCG	362
Qy	21	AlaIysAsnCysTyArMetValIlleuEnglySerserIysValGlyStrAlaIle	40
Db	363	GCCAAAGACTGCTACAGATGTCTATCTCGGCTCATCCAAAGTGGGCAAGACGGCAGTC	422
Qy	41	ValSerArgpHeleUthrGlyArgpHeGluAspaLatyrThrProthrIleGluaspHe	60
Db	423	GGTGGCGGCTTCTCAGCGCGGCTTCGAGGACCCCTTACACCCCTACCATTTGAAGACTTC	482
Qy	61	HisArglySpHetySerIleArgglyGluValtyrGluLeuaspIleleuAspThrSer	80
Db	483	CACCGAAATTTTACTCATGCTCGGGCGCAAGTCTACCAATTGGACATACTGGACATCT	542

OY	81	GlysnhisproheproAlametaArgleusertlle-----	93
Db	543	GGCATCATCCGTTTCCCCGCAATGGCGCCTCTCATCTTCACAGGTGAATGGGGACC	602
OY	93	-----	93
Db	603	GACAGGACCGTGCGGAGGGAATCTCGGGGAGCGGATGGCGGTGHTGTGCTTGG	662
OY	94	-----leu 94	
Db	663	GCTGTGCTGTCTGCTGCTCCGTCCTTGGAGCTGCCCTTCACCTTTCCACTGTTCCCTTG	722
OY	95	ThrgIyaspValPheIIeLeuValIPheserLeuAspAsnArgAspSerPheglugluVal	114
Db	723	TA-GGACAGCATTTCATCTGCTGTCTTCACGTTAGACAACCGGACATCTCTTGAGAGAGTG	781
OY	115	GlnArgLeuArggInginIleLeuAspPrhLyserCysLeuLYAsnLysThrlysGlU	134
Db	782	CAAAGGCTCAAAAGCAGATCTAGACACCAAGCTCTGTCAAGAACAACCAAAAG	841
OY	135	AsnValAspValPProleuValIIecysgIyaSnlysgIyAspArgAspPhetrArglu	154
Db	842	AATGTGACCGTCCGCTGGTCATTTGGGTAAACAAAAGSGGACCGGACTTCTACCGCGAA	901
OY	155	ValAspGlnArggluilegIngleuValGIyAspAspProGlnArgCysAlatyRhe	174
Db	902	GTGGAGCAGCGGGGAGATTGAGACAGCTGTGTGGCGGATGACCTCTCACGTTGTCCCTAC	961
OY	175	GluIleSerAlalyLyAsnSerSerLeuAspBclMetPheArgAlaleuPhealMet	194
Db	962	GAGATCTCGCCCAAGAAGATACAGACCTGGACCAATCTTCCGTCGCTCTTGGCCATG	1021
OY	195	AlalySeruProserglumetSerProAspLeuHISArgLyValserValgIntyrCys	214
Db	1022	GCCAAGCTGCTTACGAGATGAGCCCTGACCTTCACCGCAAGGTGTCTGTGCAAGTACTGT	1081
OY	215	AspValleuHISLyLySAlaleuArgAsnLyLySAlaleuArgAlagIyserglyLy	234
Db	1082	GACGTGCTGCACAAAAAGCGCTCGAGAACAAAGACTTCTGGGTGGCGGAGC---GSA	1138
OY	235	GlygllyAspProgllyAspAlapheglYIlevalAlaProPhealArgArProser	254
Db	1139	GGTGGGGCGACGACGAGATGCTTTGGCATCTTGGCGCCCTTGTCTGCGAGACTAGC	1198
OY	255	ValHisSerApPleuMetYrIleArggluLySAlaserAlaglYserGlnaIalysAsp	274
Db	1199	GTGCTATACGACCTCATGTACTTCGTGACAAAACAGTGTCAAGCAGCAGGCTAAGGAC	1258
OY	275	LysgluarGcysvalIIeser 281	
Db	1259	AAGGAGCGCTGTGCATCACTGT 1279	
RESULT 5			
; Sequence 5, Application US/08247946A			
; Patent No. 5792638			
; GENERAL INFORMATION:			
; APPLICANT: AARONSON, S.A.; CHAN, A.;			
; APPLICANT: MIKI, T.			
; TITLE OF INVENTION: NOVEL HUMAN RAS-RELATED			
; TITLE OF INVENTION: ONCOGENES UNMASKED BY EXPRESSION OF CDNA			
; NUMBER OF SEQUENCES: 11			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: MORGAN & FINNEGAN			
; STREET: 345 PARK AVENUE			
; CITY: NEW YORK			
; STATE: NEW YORK			
; COUNTRY: USA			
; ZIP: 10154			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: FLOPPY DISK			

```

COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/247,946A
FILING DATE: 24-MAY-1994
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: DOROTHY R. AUTH
REGISTRATION NUMBER: 36,434
REFERENCE/DOCKET NUMBER: 2026-4150
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792
SEQUENCE CHARACTERISTICS:
INFORMATION FOR SEQ ID NO: 5:
LENGTH: 615
TYPE: Nucleic acid
STRANDEDNESS: Double
MOLECULE TYPE: Unknown
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Human
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE:
CELL TYPE:
CELL LINE:
ORGANELLE:
FEATURE:
NAME/KEY: TC21 gene
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION:
US-08-247-946A-5

Alignment Scores:
Pred. No.: 2,45e-27 Length: 615
Score: 300.00 Matches: 71
Percent Similarity: 55.32% Conservative: 33
Best Local Similarity: 37.77% Mismatches: 70
Query Match: 20,70% Indels: 14
DB: 1 Gaps: 3

US-09-709-103-2 (1-281) x US-08-247-946A-5 (1-615)
QY 25 TYARGMETVALILELEUGLYSERLYSVALGLYSTRHRAIALLEVALSERARGPHE 44
DB 43 TACCGGCTCGTGGTGGCGGGCGGGCGGTGGCAAGTGGCGGTCCACCATCCAGTTC 102
QY 45 LEUTHRGLYARGPHEGLUASPALATYRTHRPROTHRILEGLUASPPHEHISARGLYS 64
DB 103 ATCCAGTCCATTGTTGAACGATTTATGATCCACCATTTGAATCTTACACAAAGCAG 162
QY 65 TYRSERILEARGLYGLUVALITYRGINLEUASPILEUASPTHRISERGLYASNHS 84
DB 163 TGTGTATATACATGACAGACAGCCCGCTAGATTTTGGATACACAGAGANNNGAAG 222
QY 85 PHEPRALAMETARGARGLEUSERILELEUTHRGLYASVALPHEHILEUVALPHE 104
DB 223 TTTGGAGCCATGACAGACGATATAGAGACTGGCGAAGCCTTCTGTTGGTCTTTTCA 282
QY 105 LEUASPSANARGSPSERPHEGLUGLUVALGINARGLEUARGINGINLEUASPTHR 124
DB 283 GTCACAGATAGAGCAGATTGTAAGAAATCTATAGTTTCAAGACAGATTC----- 336
QY 125 LYSERCYSLEUYSANLYSTRYSGLUASVALASPYVALPROLEUVALILECYSG 144
DB 337 -----AGAGTAAAGATCTGATGAGTGTCCCATGATTTTAATTGGT 378
```

```

QY 145 ASNLYSGLYASPARGASPHETRYARGLUVALASPGILNARGLUILEGLUGINLEUVAL 164
DB 379 AATAAAGCATCTGCATATCAACACAGTACACAGACAGACAGACAGACAGATTAGCA 438
QY 165 GLYASPSAPROGLINARGCYSALATYRPHGLUILESERALLYSLSYASNSERLEU 184
DB 439 CGGCAG-----CTTAAGTATACATACATGAGGCATCAGCAAAAGTATGATGAATGTA 492
QY 185 ASPGILMEPHEARGALAEUPHEALAMETALALYSLEUPROSERGLUJET----- 201
DB 493 GATCAAGCTTTCCATGACTGTCGCGGTATCAAGAAATTTCAAGACAGAGAAATGTCCT 552
QY 202 ---SerProAspLeuHisArgLys 208
DB 553 CTTTCAACAGAACCAACACAGGANA 576

RESULT 6
PCT-US95-06420-5
Sequence 5, Application PC/TUS9506420
GENERAL INFORMATION:
APPLICANT: AARONSON, S.A.; CHAN, A.;
APPLICANT: MIKI, T.
TITLE OF INVENTION: NOVEL HUMAN RAS-RELATED
TITLE OF INVENTION: ONCOGENES UNMASKED BY EXPRESSION OF CDNA
NUMBER OF INVENTIONS: 5
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/06420
FILING DATE: 24-MAY-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/247,946
FILING DATE: 24-MAY-1994
ATTORNEY/AGENT INFORMATION:
NAME: DOROTHY R. AUTH
REGISTRATION NUMBER: 36,434
REFERENCE/DOCKET NUMBER: 2026-4150PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 615
TYPE: Nucleic acid
STRANDEDNESS: Double
MOLECULE TYPE: Unknown
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Human
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE:
CELL TYPE:
CELL LINE:
ORGANELLE:
FEATURE:
```



```

Db      649 GATGATCCCTCTATACATTAGTTCGAGAAATTCGAAACATTAAGAAAGATGAGCANA 708
Oy      204 AspleuHisArgLys 208
Db      709 GATGTAAGAAAAGAG 723

RESULT 8
PCT-US93-06251-29
Sequence 29, Application PC/TUS9306251
GENERAL INFORMATION:
APPLICANT: Wicksom, Eric and Rife, Jason P.
TITLE OF INVENTION: Trivalent Synthesis of Oligonucleotides Containing
TITLE OF INVENTION: Stereospecific Alkylphosphonates and Arylphosphonates
NUMBER OF SEQUENCES: 93
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCULLY, SCOTT MURPHY & PRESSER
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: NY
COUNTRY: USA
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/06251
FILING DATE: 19930830
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 8586
TELECOMMUNICATION INFORMATION:
TELEPHONE: 516-742-4343
TELEFAX: 516-742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 5775 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
PCT-US93-06251-29

Alignment Scores:
Pred. No.: 2.37e-23 Length: 5775
Score: 279.00 Matches: 67
Percent Similarity: 54.59% Conservative: 34
Best Local Similarity: 36.22% Mismatches: 72
Query Match: 19.25% Indels: 12
DB: Gaps: 4

US-09-709-103-2 (1-281) x PCT-US93-06251-29 (1-5775)
Oy      25 TyrArgMetValIleLeuGlySerSerLysValGlyLysThrAlaIleValSerArgPhe 44
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      202 TATAAAGCTTGTGCTGAGAGCTGTGGCGTGAAGCAGAGTGCCTTACACATTCACAGCTA 261
Oy      45 LeuThrGlyArgPheGluAspAlaTyrThrProThrIleGluAspPheHisArgLysPhe 64
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      262 ATTCAGAAATCATTTTGTGAGCAATATGATCCACAAATAGAGATTCCTACAGAAAGCAA 321
Oy      65 TyrSerIleArgGlyGluValTyrGlnLeuAspIleLeuAspThrSerGlyAsnHisPro 84
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      322 GTAGTAATGATGAGAAACCTGCTCTTGGATATTCGACACAGCAGCTCAAGAGAG 381
Oy      85 PheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIleLeuValPheSer 104
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      382 TACAGTCGAATGAGGAGACACATGACATGAGAGACTGGGAGAGGCTTCTTGTGCTATTGCC 441

```

```

Oy      105 LeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGlnGlnIleLeuAspThr 124
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      442 ATAAATATACATAAATCATTTGAAGATATATCCATTATAGAAACAAT----- 492
Oy      125 LysSerCysLeuLysAsnLysThrIleGluAsnValPheLeuValIleCysGly 144
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      493 -----AAAAGATTAGGAGCTGTGAGAGATGACTATGCTCTTGTAGGA 537
Oy      145 AsnLysGlyAspArgAspPheTyrArgGluValAspGlnArgGluIleGluGlnLeuVal 164
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      538 AATAATGATGATTGCC---TCTAGAACAGTAGACACAAAAGAGGCTCAGACTTAGCA 594
Oy      165 GlyAspAspProGlnArgCysAlaTyrPheGluIleSerAlaLysAsnSerIleu 184
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      595 AGAAGT-----TATGAATTCCTTTATTTGAAACATCAGAAAAGACAGAGGCTGT 648
Oy      185 AspGlnMetPheArgAlaLeuPhe---AlaMetAlaLysLeuProSerGlnMetSerPro 203
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      649 GATGATGCCTTCTATACATTAGTTCGAGAAATTCGAAACATTAAGAAAGATGAGCAA 708
Oy      204 AspleuHisArgLys 208
Db      709 GATGTAAGAAAAGAG 723

RESULT 9
US-08-884-866A-2
Sequence 2, Application US/0884866A
GENERAL INFORMATION:
APPLICANT: Chien, Shu
APPLICANT: Shyu, John Y-J
TITLE OF INVENTION: GENE THERAPY IN CORONARY ANGIOPLASTY AND
TITLE OF INVENTION: BYPASS
FILE REFERENCE: UCSD1100-1
CURRENT APPLICATION NUMBER: US/08/884,866A
CURRENT FILING DATE: 1997-06-30
PRIOR APPLICATION NUMBER: 60/030,358
PRIOR FILING DATE: 1996-11-08
NUMBER OF SEQ ID NOS: 25
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 570
TYPE: DNA
ORGANISM: Rattus norvegicus
FEATURE:
NAME/KEY: CDS
LOCATION: (1)...(570)
US-08-884-866A-2

Alignment Scores:
Pred. No.: 1.18e-24 Length: 570
Score: 277.50 Matches: 73
Percent Similarity: 53.23% Conservative: 34
Best Local Similarity: 36.32% Mismatches: 72
Query Match: 19.15% Indels: 23
DB: Gaps: 6

US-09-709-103-2 (1-281) x US-08-884-866A-2 (1-570)
Oy      25 TyrArgMetValIleLeuGlySerSerLysValGlyLysThrAlaIleValSerArgPhe 44
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      10 TATAAGCTGTGAGAGCTGTGGCGGCGGCGGTGGGCAAAATGCGCTGACACATTCAGCTG 69
Oy      45 LeuThrGlyArgPheGluAspAlaTyrThrProThrIleGluAspPheHisArgLysPhe 64
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      70 ATCCAGAACCATTTTGTGAGCAATATGATCCACAAATAGAGATTCCTACAGAAAGCAA 129
Oy      65 TyrSerIleArgGlyGluValTyrGlnLeuAspIleLeuAspThrSerGlyAsnHisPro 84
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      130 GTGTCATTTAGTGGAGAGAGTGCCTTGTGAGATCTGATACCCGCGGCTGAGAGAG 189
Oy      85 PheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIleLeuValPheSer 104
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      190 TACAGCGCATGCGGAGACCAATGCAATGCGACCGGAGAGGCTTCTTGTGCTATTGCC 249

```


OY	105	LeuAspArgAspSerPheGluGlnValGIInmrgLeuArgGlnGlnIleLeuAspThr	124
Dd	250	ATCAACAACACCACCAAGTCTTTTGAGAGACACTCCACAGTACAGGGAGGCAGTC-----	300
OY	125	LysSerCysLeuLysAsnIysThrLysGluAsnValAspValProLeuValIleCysGly	144
Dd	301	-----AAAGGGGTGAAGAGACTCGGATATACGTGCCCATAGTCTCTGCTGGGG	345
OY	145	AsnLysGlyAspArgAspPheTYrArgGluValAspIleArgIuIleGlnLeuVal	164
Dd	346	AACAAGTGTGACCTG---GCTGCACGCACACTGTGGAAATCTCGGCAGGCTCAGAACCTGCC	402
OY	165	GlyAspAspProGlnIArgCysAlaTYrPheGluIleSerAlaLysLysAsnSerLeu	184
Dd	403	CGAAC-----TACGGCATTCCTCCATCATGTAGACCTCGGCCAAGCCCGCAGGAGAGTC	456
OY	185	AspGlnMetPheArgAlaLeu-----PheAlaMetAlaLysLeu---	197
Dd	457	GAGGATGGCTTCTACAGTGTGGTCCGTGAGATCCGGCAGACAAACGTGCGGAAGCTGAAC	516
OY	198	---ProSerGlnMetSerProAspLeuHisArgLysValSerValGlnTYrCysAspVal	216
Dd	517	CCCTCCTATGAGAGTGGCCCCGG-CTGCAT-----GAGGTGCAGAAGTGTGTCTCTC	566
OY	217	Leu	217
Dd	567	CTG	569

```

RESULT 10
US-08-884-866A-11
? Sequence 11. Application US/08884866A
? GENERAL INFORMATION:
? APPLICANT: Chien, Shu
? APPLICANT: Shyy, John Y-J
? TITLE OF INVENTION: GENE THERAPY IN CORONARY ANGIOPLASTY AND
? TITLE OF INVENTION: BYPASS
? FILE REFERENCE: UCSD100-1
? CURRENT APPLICATION NUMBER: US/08/884,866A
? CURRENT FILING DATE: 1997-06-30
? PRIOR APPLICATION NUMBER: 60/030,358
? PRIOR FILING DATE: 1996-11-08
? NUMBER OF SEQ ID NOS: 25
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 11
? LENGTH: 570
? TYPE: DNA
? ORGANISM: Artificial Sequence
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (49)...(51)
? OTHER INFORMATION: nm = Any nucleic acid triplet, except for UCA
? OTHER INFORMATION: UCG, UCG, UCU, AGC, or AGU
? OTHER INFORMATION: Variation of SEQ ID NO.:2
US-08-884-866A-11

```

Pred. NO.:	1,18e-24	Length:	570
Alignment Scores:		Matches:	34
Score:	277.50	Conservative:	73
Percent Similarity:	53.23%	Mismatches:	72
Best Local Similarity:	36.32%	Indels:	23
Query Match:	19.15%	Gaps:	6
DB:	4		

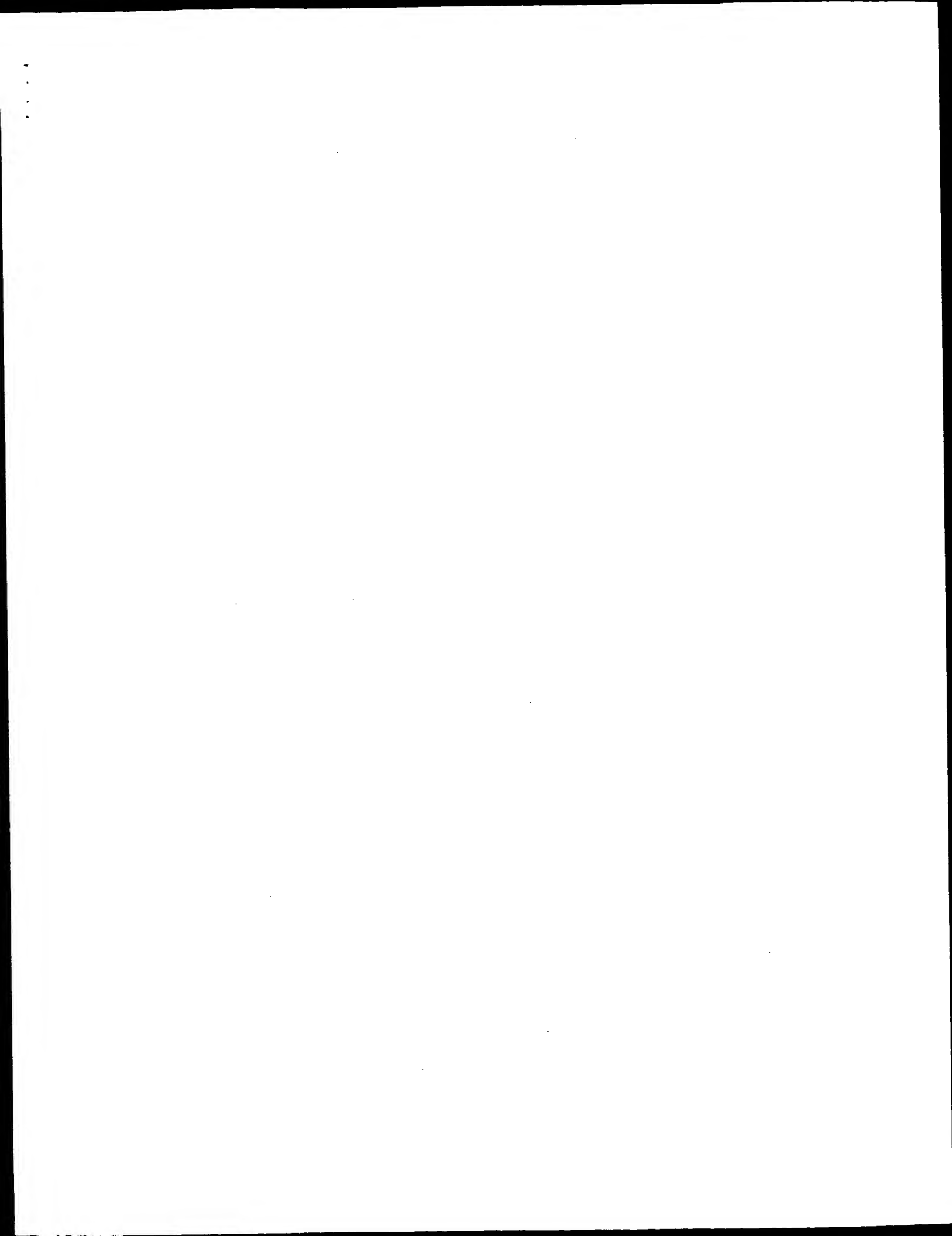
Uy YyAirmetValIleEngLIysSerIysValGIlysthrAlalValSerIphe 44
 Db TATAACTGTGATGGTGTGGCGCCGCAGTGTTGGCAAAANNNGCTGACATTCCAGCTG 69
 Oy LerrThglYArpPhgUlaspaLlaTyThrProthrILeGUaspheIsArgLySphe 64
 Db ATTCACAACCATTTTGTGGACCAATACAGCCCACCTATACAGGATTCCTCAGCGGAAGAG 120

OY		65	TyrSerIleArgGlyValIleuValTyrLeuIleuAspIleuAspThrSergLyAsnHisPro	84
Dp		130	GtGGcATtTGAtGGGGAGACgCGCTGTtTGACAATCCtGGATACCcCGCcCGCTGAGAG	189
OY		85	PheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIleLeuValPheSer	104
Dd		190	TTCACCGGCcCATCGGGACCACTCAtGCATGCCAcCGGGAgGGcCTTCtGTGTGTtTTGCC	244
OY		105	LeuAspAsnArgAspSerPheGluGluValGlnIarGleuArGtGInIleLeuAspThr	122
Dd		250	ATCACAAACACCACACTCTtTGAGGACATCCACCAAGTACAGGAGCAATC-----	300
OY		125	LysSerCySLeuLysAsnLysThrIleGslGlnAsnValAspValProLeuValIleCyGly	144
Dd		301	-----AAACGGGTGAAGACtCGGATGACtCGCCCAATGctGTCTGTCTGGCG	345
OY		145	AsnLysGlyAspArgAspPheTyrAlaGlyValAspGlnArgIuIleGluGlnLeuVal	164
Dd		346	AACAGATGTGACtCGT---GCTGCACAGCACTGTGGAAATCTCGCGAGGCTCGAGCACTCGCC	402
OY		165	GlyAspAspProGlnIarGySalatYrPheGluIleSerAlaLysLysAsnSerSerLeu	184
Dd		403	CGAAGC-----TACGGCATCCCCtACATGTGAGAACCTCGCGCAAGACCCCGCAGGAGTG	456
OY		185	AspGlnMetPheAlaGlyAlaLeu-----PheAlaMetAlaLysLeu--	197
Dd		457	GAGAGTCCTtTCACACCTGGTGGTCGtGAGATCCGGCAGCAGCAACAAGCTCGGAAGCTGAAc	516
OY		198	--ProSergIuMetSerProAspLeuHISArgLySvalSerValGlnTyrCyAspAryal	216
Dd		517	cCTCTCATGTAGAGTAGGCCCGCG- CTGCAT-----GAGCTGCACAAAGTGTGTGCTCTC	566
OY		217	Leu	217
Dd		567	CTG 569	

RESULT 11
US-08-429-964-85
Sequence 85, Application US/08429964
Patent No. 5962243
GENERAL INFORMATION:
APPLICANT: BROWN, MICHAEL S.
APPLICANT: GOLDSTEIN, JOSEPH L.
APPLICANT: REISS, YUVAL
APPLICANT: JAMES, GUY L.
TITLE OF INVENTION: METHODS FOR THE IDENTIFICATION OF FARNESYL
TITLE OF INVENTION: TRANSFERASE INHIBITORS
NUMBER OF SEQUENCES: 85
CORRESPONDENCE ADDRESS:
ADDRESSEE: ARNOLD, WHITE & DUKKEE
STREET: P.O. BOX 4433
CITY: HOUSTON
STATE: TEXAS
COUNTRY: UNITED STATES OF AMERICA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/429,964
FILING DATE: 27-APR-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/021,625
FILING DATE: 16-FEB-1993
CLASSIFICATION: 435
APPLICATION NUMBER: US 07/822,011
FILING DATE: ABANDONED
CLASSIFICATION: 435
APPLICATION NUMBER: PCT/US/91/02650

RESULT 15
US-08-429-964-83
; Sequence 83, Application US/08429964
; Patent No. 5962243
GENERAL INFORMATION:
APPLICANT: BROWN, MICHAEL S.
APPLICANT: GOLDSTEIN, JOSEPH L.
APPLICANT: REISS, YUVAL
APPLICANT: JAMES, GUY L.
TITLE OF INVENTION: METHODS FOR THE IDENTIFICATION OF FARNESYL
TITLE OF INVENTION: TRANSFERASE INHIBITORS
NUMBER OF SEQUENCES: 85
CORRESPONDENCE ADDRESS:
ADDRESSEE: ARNOLD, WHITE & DURKEE
STREET: P.O. BOX 4433
CITY: HOUSTON
STATE: TEXAS
COUNTRY: UNITED STATES OF AMERICA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/429,964
FILING DATE: 27-APR-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/021,625
FILING DATE: 16-FEB-1993
CLASSIFICATION: 435
APPLICATION NUMBER: US 07/822,011
FILING DATE: ABANDONED
CLASSIFICATION: 435
APPLICATION NUMBER: PCT/US/91/02650
FILING DATE: 18-APR-1991
CLASSIFICATION: 435
APPLICATION NUMBER: US 07/615,715
FILING DATE: 20-NOV-1990
CLASSIFICATION: 435
APPLICATION NUMBER: US 07/510,706
FILING DATE: 18-APR-1990 (ABANDONED)
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: PARKER, DAVID L.
REGISTRATION NUMBER: 32,165
REFERENCE/DOCKET NUMBER: USPD-432/PAR
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 769-2679
TELEX: 79-0924

Job time : 87 secs



score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

(without alignments)
11465.267 Million cell updates/sec

11455.267 Million cell updates/sec

.....agcgctgcgtcatcagctag 846

residues: 4109280

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	846	100.0	946	9	AF498923	AF498923 Homo sapi
2	846	100.0	979	9	AF172846	AF172846 Homo sapi
3	846	100.0	1740	9	AF069506	AF069506 Homo sapi
4	846	100.0	1758	9	BC018041	BC018041 Homo sapi
5	838	99.1	1486	9	AF153192	AF153192 Homo sapi
6	835	98.7	1146	9	AF177335	AF177335 Homo sapi
7	678	80.1	1612	10	BC034166	BC034166 Mus musc
8	678	80.1	1516	10	AF239157	AF239157 Rattus nor
9	676.4	80.0	1633	10	AF009246	AF009246 Rattus nor
10	625	73.9	4990	9	AF228279	AF228279 Homo sapi
11	625	73.9	183334	9	AC020558	AC020558 Homo sapi
12	623.4	73.7	5141	9	AF262018	AF262018 Homo sapi
13	609.4	72.0	58886	2	AC073621	AC073621 Homo sapi
14	607.8	71.8	183598	2	AC093068	AC093068 Homo sapi
15	515	60.9	179124	10	AL603710	AL603710 Mouse DNA
16	438.4	51.8	162504	2	AC128935	AC128935 Rattus no
17	407.4	48.2	2832	6	AX393344	AX393344 Sequence
18	407.4	48.2	2832	6	AF279143	AF279143 Homo sapi
19	407.4	48.2	2973	6	AX393677	AX393677 Sequence
20	407.4	48.2	3058	9	BC013419	BC013419 Homo sapi
21	362.2	42.8	2699	9	HSB8031172	HSB8031172 Homo sapi
22	359.6	42.5	3469	10	AF134409	AF134409 Rattus no
23	358	42.3	3020	6	AX393362	AX393362 Sequence
24	240.2	28.4	114771	10	HS569919	HS569919 Human DNA
25	204.4	24.2	211071	10	AC076974	AC076974 Mus muscu
26	147.2	17.4	66631	2	AC118046	AC118046 Mus muscu
27	136.2	16.1	7774	9	AB076888	AB076888 Homo sapi
28	136.2	16.1	2827	9	AK096600	AK096600 Homo sapi
29	136.2	16.1	3391	9	BC030660	BC030660 Homo sapi
30	136.2	16.1	177540	9	AC006538	AC006538 Homo sapi
31	134.8	15.9	597	9	AY050637	AY050637 Homo sapi
32	134.8	15.9	597	9	AY050641	AY050641 Homo sapi
33	132	15.6	3061	6	AX430418	AX430418 Sequence
34	127.6	15.1	2294	10	BC026377	BC026377 Mus muscu
35	125.8	14.9	1249	6	AX430295	AX430295 Sequence
36	116	13.7	4412	9	AB062937	AB062937 Macaca fa
37	113.6	13.4	2145	9	HSB803027	HSB803027 Homo sapi
38	113.4	13.4	1913	9	BC012362	BC012362 Homo sapi
39	113.4	13.4	147994	2	AC069415	AC069415 Homo sapi
40	113.4	13.4	165793	2	AC069072	AC069072 Homo sapi
41	113.4	13.4	198170	9	AC117394	AC117394 Homo sapi
42	112.8	13.3	1875	9	BC008065	BC008065 Homo sapi
43	112.8	13.3	2505	9	AB076889	AB076889 Homo sapi
44	112.8	13.3	4167	6	AX477382	AX477382 Sequence
45	112.8	13.3	143299	9	AL353619	AL353619 Human DNA

ALIGNMENTS

RESULT 1	AF498923	846 bp	MRNA	linear	PRI 01-MAY-2002
LOCUS	AF498923				
DEFINITION	Homo sapiens activator of G protein signaling (RASD1) mRNA,				
ACCESSION	AF498923				
VERSION	AF498923				
KEYWORDS	AF498923.1	GI:20379021			
SOURCE	Homo sapiens.				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE	1 (bases 1 to 846)				
AUTHORS	Puhl,H.L., Ikeda,S.R. and Aronstam,R.S.				
TITLE	Direct Submision				

JOURNAL Submitted (05-Apr-2002) cDNA Resource Center, Guthrie Research Institute, One Guthrie Square, Sayre, PA 16840, USA

FEATURES
source
1..846
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="Brain"

gene

CDS

1..846
/gene="RASD1"
/note="AgS1"
1..846
/gene="RASD1"
/codon_start=1
/product="activator of G protein signaling"
/protein_id="AA021071.1"
/db_xref="GI:20379022"

/translation="MKLAAMIKKMPDSELSIPAKNCYRMVITLSSVGTATYSRF
LIGREDATPTIEDPHKRFYSINGEVYQDLIDSGNHPEPAMRRSLITGDEFLV
FSLNDRDSFEVYQRLQOILDTKSLKNTKENVDPVLYCGNKGDREYREVDREI
BOLVDDDPORCAVFEISAKNSLIDQMRALFPAKALPSEMSPDHRRVSVQCDVLA
KKAIRNKILIRAGSGGGGDPDAGIYAPFAPRRSVSHDLMTIREKASAGSQADKE
RCVIS"

BASE COUNT 176 a 288 c 252 g 130 t
ORIGIN

Query Match 100.0%; Score 846; DB 9; Length 846;
Best Local Similarity 100.0%; Pred. No. 1.1e-103; Mismatches 0; Indels 0; Gaps 0;
Matches 846; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGAAGTGGCCGGATGATCAAGAAGATGTGCCCGAGCGACTGGAGCTGACTATCCG 60
1 ATGAAGTGGCCGGATGATCAAGAAGATGTGCCCGAGCGACTGGAGCTGACTATCCG 60
61 GCCAAGACGCTATGTCGATGTCATCCCTGCTGCTCCAGAGTGGGCAAGAGCGCCATC 120
61 GCCAAGACGCTATGTCGATGTCATCCCTGCTGCTCCAGAGTGGGCAAGAGCGCCATC 120
121 GTGCGCGCTTCTCAGCGCGCGCTTCTGAGAGCGCTTACAGCGCTTACATCGAGGACTTC 180
121 GTGCGCGCTTCTCAGCGCGCGCTTCTGAGAGCGCTTACAGCGCTTACATCGAGGACTTC 180
121 GTGCGCGCTTCTCAGCGCGCGCTTCTGAGAGCGCTTACAGCGCTTACATCGAGGACTTC 180
181 CACCGCAAGTTTACTCCATCCCGCGGAGGTCTACAGCTCGACATCTCGACAGCTCC 240
181 CACCGCAAGTTTACTCCATCCCGCGGAGGTCTACAGCTCGACATCTCGACAGCTCC 240
241 GGCACCAACCCGTCGCCGCGATCGCGGCTCTCATCTCCACAGAGAGAGAGCTTTTCATC 300
241 GGCACCAACCCGTCGCCGCGATCGCGGCTCTCATCTCCACAGAGAGAGAGCTTTTCATC 300
241 GGCACCAACCCGTCGCCGCGATCGCGGCTCTCATCTCCACAGAGAGAGAGCTTTTCATC 300
301 CTGGGTTCAGTCTGACAAACCGGACTCTTCTGAGAGGTGACGCGCTCAGGACAGCAG 360
301 CTGGGTTCAGTCTGACAAACCGGACTCTTCTGAGAGGTGACGCGCTCAGGACAGCAG 360
361 ATCTCTGACACCAAGTCTGCTTCAAGAAACAAACAAAGAGAGAGAGAGAGAGAGAGAG 420
361 ATCTCTGACACCAAGTCTGCTTCAAGAAACAAACAAAGAGAGAGAGAGAGAGAGAGAG 420
421 GTCAATCTCGGCAACAAGAGGTGACCGGACTTCTACCGAGAGGTGAGCAACGCGAGATC 480
421 GTCAATCTCGGCAACAAGAGGTGACCGGACTTCTACCGAGAGGTGAGCAACGCGAGATC 480
421 GTCAATCTCGGCAACAAGAGGTGACCGGACTTCTACCGAGAGGTGAGCAACGCGAGATC 480
481 GAGCAGCTGTGGGGAGAGACCCCGAGCGCTTACTTCTGAGAGATCTCGGCCAAGAG 540
481 GAGCAGCTGTGGGGAGAGACCCCGAGCGCTTACTTCTGAGAGATCTCGGCCAAGAG 540
481 GAGCAGCTGTGGGGAGAGACCCCGAGCGCTTACTTCTGAGAGATCTCGGCCAAGAG 540
541 AACAGCAGCTGTGGGGAGAGACCCCGAGCGCTTACTTCTGAGAGATCTCGGCCAAGAG 600
541 AACAGCAGCTGTGGGGAGAGACCCCGAGCGCTTACTTCTGAGAGATCTCGGCCAAGAG 600
601 ATGAGCCGAGACCTGACCGAGAGGTCTCGGTGAGTACTGAGAGGTCTGACACAAAGAG 660
601 ATGAGCCGAGACCTGACCGAGAGGTCTCGGTGAGTACTGAGAGGTCTGACACAAAGAG 660

661 GCCTGCGGAAACAAGAGCTGCTGCGGCGCGAGCGCGCGCGCGCGCGAGACCCGGC 720
661 GCCTGCGGAAACAAGAGCTGCTGCGGCGCGAGCGCGCGCGCGCGCGAGACCCGGC 720
721 GAGCGCTTGGCATCGTGGACACCTTGGCGCGCGCGCGCGCGCGCGCGAGAGCTTCATG 780
721 GAGCGCTTGGCATCGTGGACACCTTGGCGCGCGCGCGCGCGCGCGCGAGAGCTTCATG 780
781 TACATTCGCGAGAGAGCGCGCGCGCGAGCGCGCAAGAGCAAGAGAGAGAGAGAGAGAG 840
781 TACATTCGCGAGAGAGCGCGCGCGCGAGCGCGCAAGAGCAAGAGAGAGAGAGAGAGAG 840
841 AGCTAG 846
841 AGCTAG 846
841 AGCTAG 846

RESULT 2
AF172846 979 bp mRNA linear PRI 05-JAN-2000
LOCUS Homo sapiens ras-related protein (DEXRAS1) mRNA, complete cds.
DEFINITION AF172846
VERSION AF172846.1 GI:6014488
KEYWORDS
SOURCE
ORGANISM
Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 979)
Tu Y. and Wu C.
Cloning, expression and characterization of a novel human
Ras-related protein that is regulated by glucocorticoid hormone
2 (bases 1 to 979)
Tu Y. and Wu C.
Direct Submission
Submitted (27-JUL-1999) Department of Cell Biology, University of
Alabama at Birmingham, 1670 University Blvd., Birmingham, AL 35294,
USA

FEATURES
source

1..979
/organism="Homo sapiens"
/db_xref="taxon:9606"
1..979
/gene="DEXRAS1"
64..909
/note="similar to Mus musculus ras-related protein encoded
by GenBank Accession Number AF009246; regulated by
glucocorticoid hormone"
/codon_start=1
/product="ras-related protein"
/protein_id="AA01364.1"
/db_xref="GI:6014489"
/translation="MKLAAMIKKMPDSELSIPAKNCYRMVITLSSVGTATYSRF
LIGREDATPTIEDPHKRFYSINGEVYQDLIDSGNHPEPAMRRSLITGDEFLV
FSLNDRDSFEVYQRLQOILDTKSLKNTKENVDPVLYCGNKGDREYREVDREI
BOLVDDDPORCAVFEISAKNSLIDQMRALFPAKALPSEMSPDHRRVSVQCDVLA
KKAIRNKILIRAGSGGGGDPDAGIYAPFAPRRSVSHDLMTIREKASAGSQADKE
RCVIS"

BASE COUNT 195 a 347 c 284 g 153 t
ORIGIN

Query Match 100.0%; Score 846; DB 9; Length 979;
Best Local Similarity 100.0%; Pred. No. 1.1e-103; Mismatches 0; Indels 0; Gaps 0;
Matches 846; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGAAGTGGCCGGATGATCAAGAAGATGTGCCCGAGCGACTGGAGCTGACTATCCG 60
64 ATGAAGTGGCCGGATGATCAAGAAGATGTGCCCGAGCGACTGGAGCTGACTATCCG 123


```

QY 61 GCCAAGAACTGCTATCCGATGTCATCTCTGCGCTGCTCCAAAGTGCGGCAAGAGCCATC 120
    |||||||
Db 124 GCCAAGAACTGCTATCCGATGTCATCTCTGCGCTGCTCCAAAGTGCGGCAAGAGCCATC 183
QY 121 GTGTGGGCTTCTCACCAGGCGCTTGAGAGAGCTACAGGCTTACATCGAGGACTTC 180
    |||||||
Db 184 GTGTGGGCTTCTCACCAGGCGCTTGAGAGAGCTTACAGGCTTACATCGAGGACTTC 243
QY 181 CACCGCAAGTTCTACTCCATCCGCGGAGGAGTTCAGAGCTCGACATCTCGACAGCTCC 240
    |||||||
Db 244 CACCGCAAGTTCTACTCCATCCGCGGAGGAGTTCAGAGCTCGACATCTCGACAGCTCC 303
QY 241 GGCACCAACCGCTTCCCGCGCATCCGCGCTCTCCATCTCCACAGAGACGTTTTCATC 300
    |||||||
Db 304 GGCACCAACCGCTTCCCGCGCATCCGCGCTCTCCATCTCCACAGAGACGTTTTCATC 363
QY 301 CTGGTGTTCAGTCTGACACACCGGAGTCTCTGAGAGAGTGCAGGCTCAGGACAG 360
    |||||||
Db 364 CTGGTGTTCAGTCTGACACACCGGAGTCTCTGAGAGAGTGCAGGCTCAGGACAG 423
QY 361 ATCTCTGACACCAAGTTTCTGCTTCAGACAAACCAAGAGAGACGTGACGTGCCCTG 420
    |||||||
Db 424 ATCTCTGACACCAAGTTTCTGCTTCAGACAAACCAAGAGAGACGTGACGTGCCCTG 483
QY 421 GTGATCTGCGGCAACAGGTTGACCGGAGTCTTACCGGAGGAGTGCAGCGCGAGATC 480
    |||||||
Db 484 GTGATCTGCGGCAACAGGTTGACCGGAGTCTTACCGGAGGAGTGCAGCGCGAGATC 543
QY 481 GAGCAGCTGTGTGGGCGACGACCCGACGCGTGGCGCTTACTTCCAGATCTGCGGCAAGAG 540
    |||||||
Db 544 GAGCAGCTGTGTGGGCGACGACCCGACGCGTGGCGCTTACTTCCAGATCTGCGGCAAGAG 603
QY 541 AACAGCAGCTGTGACGACGATGTTCCGCGGCTCTTGGCCATGGCCAAAGCTGCCAGGAG 600
    |||||||
Db 604 AACAGCAGCTGTGACGACGATGTTCCGCGGCTCTTGGCCATGGCCAAAGCTGCCAGGAG 663
QY 601 ATGAGCCCAAGCTGACGCGCAAGGTTCTGCTGAGTCTGACGACGTCGTCACCAAGAG 660
    |||||||
Db 664 ATGAGCCCAAGCTGACGCGCAAGGTTCTGCTGAGTCTGACGACGTCGTCACCAAGAG 723
QY 661 GCGCTGGGCAACAAAGAGCTGTGCGGGCGGCGAGCGGCGGCGGCGGCGGCGGCGG 720
    |||||||
Db 724 GCGCTGGGCAACAAAGAGCTGTGCGGGCGGCGAGCGGCGGCGGCGGCGGCGGCGG 783
QY 721 GAGGCTTTGGCATCGTGGGAGCCCTTTCGCGCGCGGCGGCGGCGGCGGCGGCGGCGG 780
    |||||||
Db 784 GAGGCTTTGGCATCGTGGGAGCCCTTTCGCGCGCGGCGGCGGCGGCGGCGGCGGCGG 843
QY 781 TACATCCGAGAGAGCGGACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 840
    |||||||
Db 844 TACATCCGAGAGAGCGGACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 903
QY 841 AGCTAG 846
    |||||||
Db 904 AGCTAG 909

```

```

RESULT 3
AF069506 1740 bp mRNA linear PRI 10-JUN-2000
LOCUS Homo sapiens activator of G protein signaling (AGS1) mRNA, complete
DEFINITION cds.
ACCESSION AF069506
VERSION AF069506.1 GI:4959037
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 1740)
AUTHORS Cismowski, M.J., Takesono, A., Ma, C., Lizano, J.S., Xie, X.,
Fuenkranz, H., Lanier, S.M., and Duzic, E.
Genetic screens in yeast to identify mammalian nonreceptor

```

```

FEATURES
    source
        1..1740
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /tissue_type="liver"
            /dev_stage="adult"
            /gene="AGS1"
            /gene="AGS1"
            /note="ras-related G protein: similar to Mus musculus
product encoded by GenBank Accession Number AF009246"
            /codon_start=1
            /product="activator of G protein signaling"
            /protein_id="A034206.1"
            /db_xref="GI:4959038"
            /translation="MKLAAMIKKMPDSESLIPAKNRYMVLIGSSKVGKTAIVSEF
LNGREDVATPTIEDHFKFYSIGEVYQDLIDTSGNHPAPARRSLITGDVFTIV
ESLDNRDSFEVQRIKRODIDTFSKLNKTEENDVPLVTCNGNDYREVQRI
EOLVDDPQRCATVEFISAKNSLDDQRALEFAMAKLPSEMSPLHRSVQCDVIL
KKALNNKLLRAGSGGGDPDAFGIYAPARRPSVSDIMYIRERASGSAKDKE
RCVIS"
BASE COUNT 422 a 546 c 483 g 289 t
ORIGIN
Query Match 100.0%; Score 846; DB 9; Length 1740;
Best Local Similarity 100.0%; Pred. No. 9.8e-104;
Matches 846; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAAGACTGGCGCGGATGATCAAGAAAGATGTCGCGGAGCTCGAGCTGATCCCG 60
    |||||||
Db 146 ATGAAGACTGGCGCGGATGATCAAGAAAGATGTCGCGGAGCTCGAGCTGATCCCG 205
QY 61 GCCAAGAACTGCTATCCGATGTCATCTCTGCGCTGCTCCAAAGTGCGGCAAGAGCCATC 120
    |||||||
Db 206 GCCAAGAACTGCTATCCGATGTCATCTCTGCGCTGCTCCAAAGTGCGGCAAGAGCCATC 265
QY 121 GTGTGGGCTTCTCACCAGGCGCTTGAGAGAGCTACAGGCTTACATCGAGGACTTC 180
    |||||||
Db 266 GTGTGGGCTTCTCACCAGGCGCTTGAGAGAGCTTACAGGCTTACATCGAGGACTTC 325
QY 181 CACCGCAAGTTCTACTCCATCCGCGGAGGAGTCTACAGCTCGACATCTCGACAGCTTC 240
    |||||||
Db 326 CACCGCAAGTTCTACTCCATCCGCGGAGGAGTCTACAGCTCGACATCTCGACAGCTTC 385
QY 241 GGCACCAACCGCTTCCCGCGCATCCGCGCTCTCCATCTCCACAGAGACGTTTTCATC 300
    |||||||
Db 386 GGCACCAACCGCTTCCCGCGCATCCGCGCTCTCCATCTCCACAGAGACGTTTTCATC 445
QY 301 CTGGTGTTCAGTCTGACACACCGGAGTCTCTGAGAGAGTGCAGGCTCAGGACAG 360
    |||||||
Db 446 CTGGTGTTCAGTCTGACACACCGGAGTCTCTGAGAGAGTGCAGGCTCAGGACAG 505
QY 361 ATCTCTGACACCAAGTTTCTGCTTCAGACAAACCAAGAGAGACGTGACGTGCCCTG 420
    |||||||
Db 506 ATCTCTGACACCAAGTTTCTGCTTCAGACAAACCAAGAGAGACGTGACGTGCCCTG 565
QY 421 GTGATCTGCGGCAACAGGTTGACCGGAGTCTTACGCGGAGTGCAGGCTCAGGACAG 480
    |||||||
Db 566 GTGATCTGCGGCAACAGGTTGACCGGAGTCTTACGCGGAGTGCAGGCTCAGGACAG 625
QY 481 GAGCAGCTGTGTGGGCGACGACCCGACGCGTGGCGCTTACTTCCAGATCTGCGGCAAGAG 540

```

```

Db 626 GAGCAGCTGTGGCGAGACCCAGCGCTGACCTTCTGAGATCTGGCCAAAG 685
QY 541 AACAGCAGCCGTGACCAAGATGTTCCGGCGCTCTTCCGCAATGGCCAGCGCCAGGAG 600
Db 686 AACAGCAGCCGTGACCAAGATGTTCCGGCGCTCTTCCGCAATGGCCAGCGAG 745
QY 601 ATGAGCCAGACCTGACCAAGCTGTCGTCAGTACACGACGTCGTCAGCAAGAG 660
Db 746 ATGAGCCAGACCTGACCAAGCTGTCGTCAGTACACGACGTCGTCAGCAAGAG 805
QY 661 GCGCTGGGGAACAAGAGCTGCTGCGGCGCGGACGCGCGCGCGCGCGCGCG 720
Db 806 GCGCTGGGGAACAAGAGCTGCTGCGGCGCGGACGCGCGCGCGCGCGCGCG 865
QY 721 GAGCGCTTGGCATGTCGCGACCTTGGCGCGCGCGCGCGCGCGCGCGCGCATG 780
Db 866 GAGCGCTTGGCATGTCGCGACCTTGGCGCGCGCGCGCGCGCGCGCGCGCATG 925
QY 781 TACATCCGCGAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCATC 840
Db 926 TACATCCGCGAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCATC 985
QY 841 AGCTAG 846
Db 986 AGCTAG 991

```

```

RESULT 4
LOCUS BC018041 1758 bp mRNA linear PRI 06-DEC-2001
DEFINITION Homo sapiens, RAS, dexamethasone-induced 1, clone MGC:26290
IMAGE:4797482, mRNA, complete cds.
ACCESSION BC018041
VERSION BC018041.1 GI:17390075
KEYWORDS MGC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1758)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (03-DEC-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www.shgc.stanford.edu
Contact: (Dickson, Mark) mcdgexil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

```

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
 Series: IRAC Plate: 32 Row: f Column: 11.

FEATURES

SOURCE

```

1..1758
/organism="Homo sapiens"
/db_xref="LocusID:51655"
/db_xref="taxon:9606"
/clone="MGC:26290 IMAGE:4797482"
/tissue_type="Brain, hypothalamus"
/clone_lib="NIH_MGC_96"
/lab_host="DH10B"

```

```

CDS
/note="Vector: pBluescript"
213..1058
/codon_start=1
/product="RAS, dexamethasone-induced 1"
/protein_id="AAH18041.1"
/db_xref="GI:17390076"
/translation="MKLAAMIKKCPDSELSIPAKNCYRMVILGSSKVGKTAIVSR
LGRREDAYTPPIEDFHRKPYFSIREVYQDILDTSGNHPEPARRLSLITGVDFILV
FSIDNRDSFEVQRLOQILDTKSLKTKENYDPLVIGKNGDRDFEVDQREI
EQLVDDPDRCAVEFELISAKNSLIDOMRALEFMAAKLPESEPLHRRVSYQCDVLH
KKALRNKLLRAGSGGGGDPDAFVAPFARRRPSVSLMTIREASASQAKDKE
KVIS"
BASE COUNT 387 a 576 c 501 g 294 t
ORIGIN
Query Match 100.0%; Score 846; DB 9; Length 1758;
Best Local Similarity 100.0%; Pred. No. 9.8e-104;
Matches 846; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAACCTGGCGCGATGATCAAGAAGATGTGCCGAGCTCGAGCTGATATCCG 60
Db 213 ATGAACCTGGCGCGATGATCAAGAAGATGTGCCGAGCTCGAGCTGATATCCG 272
QY 61 GCCAAGACTGCTATGCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
Db 273 GCCAAGACTGCTATGCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 332
QY 121 GTGTGGGCTTCTTCAACCGCGCGCTTCAAGAGCGCTTCAAGCGCTTCAAGCGCTT 180
Db 333 GTGTGGGCTTCTTCAACCGCGCGCTTCAAGAGCGCTTCAAGCGCTTCAAGCGCTT 392
QY 181 CACCGAAGTTTACATCCGCGCGAGGCTTACACCTGACATCCCTGACACGCTTC 240
Db 393 CACCGAAGTTTACATCCGCGCGAGGCTTACACCTGACATCCCTGACACGCTTC 452
QY 241 GGCACACACCGCTTCCCGCATGCGCGCGCTTCTCATCTTCAACAGAGAGCTTTTCATC 300
Db 453 GGCACACACCGCTTCCCGCATGCGCGCGCTTCTCATCTTCAACAGAGAGCTTTTCATC 512
QY 301 CTGGTGTTCAGTTGGACAACCGCGACTCTTCAAGAGAGTGCAGGCGCTCAGGAGCAG 360
Db 513 CTGGTGTTCAGTTGGACAACCGCGACTCTTCAAGAGAGTGCAGGCGCTCAGGAGCAG 572
QY 573 ATCCTGACACCAAGTCTTCCCTCAAGAAACCAACCAAGAGAGTGCAGTGCCTG 632
QY 421 GTCATCTGGCGCAACAAGGTGACCGGCGACTTCTACCGGAGTGCAGGCGCGAGATC 480
Db 633 GTCATCTGGCGCAACAAGGTGACCGGCGACTTCTACCGGAGTGCAGGCGCGAGATC 692
QY 481 GACACAGCTGTGGGCGACGACCCCGAGCGCTGCGCTTCTGAGATCTCGGCGCAAGAAG 540
Db 693 GACACAGCTGTGGGCGACGACCCCGAGCGCTGCGCTTCTGAGATCTCGGCGCAAGAAG 752
QY 541 AACAGCAGCTGACCAAGATGTTCCGGCGCTTCTGCGCATGAGCCAGCTGCCAGCGAG 600
Db 753 AACAGCAGCTGACCAAGATGTTCCGGCGCTTCTGCGCATGAGCCAGCTGCCAGCGAG 812
QY 601 ATGAGCCAGACCTGACCAAGCTGTCGTCAGTACACGACGTCGTCAGCAAGAG 660
Db 813 ATGAGCCAGACCTGACCAAGCTGTCGTCAGTACACGACGTCGTCAGCAAGAG 872
QY 661 GCGCTGGGGAACAAGAGCTGTCGCGGCGCGGACGCGCGCGCGCGCGCGCGCG 720
Db 873 GCGCTGGGGAACAAGAGCTGTCGCGGCGCGGACGCGCGCGCGCGCGCGCGCG 932
QY 721 GAGCGCTTGGCATGTCGCGACCTTGGCGCGCGCGCGCGCGCGCGCGCGCATG 780
Db 933 GAGCGCTTGGCATGTCGCGACCTTGGCGCGCGCGCGCGCGCGCGCGCGCATG 992
QY 781 TACATCCGCGAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCATC 840

```

DB	993	TACATCCGGAGAACGCCAGCCGGCGAGCCAGCGCCAGACAGACAGGAGACGCGTGCATC	1052
QY	841	AGCTAG 846	
Db	1053	AGCTAG 1058	
RESULT 5			
AF153192			
LOCUS	AF153192	1187 bp	mRNA linear PRI 02-JUN-1999
DEFINITION	Homo sapiens ras-related protein mRNA, complete cds.		
ACCESSION	AF153192		
VERSION	AF153192.1	GI:4960166	
KEYWORDS			
SOURCE			
ORGANISM	Homo sapiens.		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (bases 1 to 1187)		
JOURNAL	Kempainen, R.J.		
REFERENCE	Identification of human pituitary Dextral		
AUTHORS	2 (bases 1 to 1187)		
TITLE	Unpublished		
JOURNAL	Kempainen, R.J.		
REFERENCE	Direct Submission		
AUTHORS	Submitted (20-May-1999) Anatomy & Physiology, Auburn University,		
JOURNAL	College of Veterinary Medicine, Auburn, AL 36849, USA		
FEATURES			
source	1..1187		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/chromosome="2"		
	/map="2q32"		
	/tissue-type="pituitary gland"		
	1..846		
	/note="Dextral: similar to Mus musculus ras-related protein encoded by GenBank Accession Number AF009246"		
	/codon_start=1		
	/product="ras-related protein"		
	/protein_id="AAD34621.1"		
	/db_xref="GI:4960167"		
	/translation="MKLAAMTKMKPSDSLELTPAKNCRYVILGSSKVGATATVSRFLTGREDATVPTIEDFRHKRYSIRGEVYQDILDTSGNHPPARRRSLIGDVFILVPSLDNRDSFEVQRLRQOILLDTKSKLKNKTEVDVPLVTCGNKGDHFEVQRELEQALGDDPQRCAYFEISAKNSSLQDMFRAFLPAKLPSEMSPLHRRKVSQYCDVLHKKALRNKKLLRAGSGGGGPGDAFGIVAFARRPSVSHDLMYIREKASAGSQAKDKREKCVTS"		
	RCVTS"		
BASE COUNT	249 a 396 c 362 g 180 t		
ORIGIN			
Query Match	99.1%; Score 838; DB 9; Length 1187;		
Best Local Similarity	99.4%; Pred. No. 1,2e-102;		
Matches 841; Conservative	0; Mismatches 5; Indels 0; Gaps 0;		
QY	1	ATGAACACGCGCGCATATCAAGAAGATGTCGCCGAGCGAGCTCGAGCTGAGTATCCG	60
Db	1	ATGAACACGCGCGCATATCAAGAAGATGTCGCCGAGCGAGCTCGAGCTGAGTATCCG	60
QY	61	GCCAAAGACTCTATCGATCGATGTCATCCTCGGCTCGTCCAAAGTGGGGAAGCGGCATC	120
Db	61	GCCAAAGACTCTATCGATCGATGTCATCCTCGGCTCGTCCAAAGTGGGGAAGCGGCATC	120
QY	121	GTCGCGCTTCCTCACCGCGCGCTTCGAGAGACGCGCTACAGCGCTTACCATCGAGACTTC	180
Db	121	GTCGCGCTTCCTCACCGCGCGCTTCGAGAGACGCGCTTACCATCGAGACTTC	180
QY	181	CACCGCAAGTCTACTCCATCGCGGCGCAGGTCTACCAAGCTCGACATCCTCGACACGTC	240
Db	181	CACCGCAAGTCTACTCCATCGCGGCGCAGGTCTACCAAGCTCGACATCCTCGACACGTC	240
QY	241	GGAACACACCGCTTCGCGGCGCATGCGGCGCTTCGATCACTTCACAGAGGAGAGCTTTTCAC	300
Db	241	GGAACACACCGCTTCGCGGCGCATGCGGCGCTTCGATCACTTCACAGAGGAGAGCTTTTCAC	300

QY	301	CTGGGTTTCAGCTGGGACAAACCGGCACTCTTCGAGGAGGTGCACGGCTCAGGACGAG	360
Db	301	CTGGTGTTCAGCTTGTGGACAAACCGGCACTCTTCGAGGAGGTGCACGGCTCAGGACGAG <td>360</td>	360
QY	361	ATCCCTGACACCAACTTCTTGCTCAAGAAACAAACCAAGAGAGAGAGTGGAGCTGCCCTG	420
Db	361	ATCCCTGACACCAACTTCTTGCTCAAGAAACAAACCAAGAGAGAGAGTGGAGCTGCCCTG	420
QY	421	GTCAATCTGGCGGCAACAAGGTTGACCGGCACTTCTACCGCGAGGTGGACAGCGAGATTC	480
Db	421	GTCAATCTGGCGGCAACAAGGTTGACCGGCACTTCTACCGCGAGGTGGACAGCGAGATTC	480
QY	481	GAGCAGCTGTGGTGGGAGAGACCCCGAGGGCTGGCTACTTGTGAATGTCGGGCGCAAGAG	540
Db	481	GAGCAGCTGTGGTGGGAGAGACCCCGAGGGCTGGCTACTTGTGAATGTCGGGCGCAAGAG	540
QY	541	AACACAGACCTGTGGACCAAGTGTCCGCGCGCTCTTCTGCGCATGGCCAGCTGCCACGAG	600
Db	541	AACACAGACCTGTGGACCAAGTGTCCGCGCGCTCTTCTGCGCATGGCCAGCTGCCACGAG	600
QY	601	ATGACACCAAGACCTTCACCGCAAGTCTTCGTTGCTGACAGTCTGCGAGCTGTCACAAAG	660
Db	601	ATGACACCAAGACCTTCACCGCAAGTCTTCGTTGCTGACAGTCTGCGAGCTGTCACAAAG	660
QY	661	GCGGTCGCGGAACAAAGAGTGTGCTGGGGCGCGGACCGGGGGGGGGGGGCGGACCCGGGC	720
Db	661	GCGGTCGCGGAACAAAGAGTGTGCTGGGGCGCGGACCGGGGGGGGGGGGCGGACCCGGGC	720
QY	721	GACGCTTTGGCATGTGGCACTCTTCGCGCGCCCGCCAGCGTACAGACGACCTCTCATG	780
Db	721	GACGCTTTGGCATGTGGCACTCTTCGCGCGCCCGCCAGCGTACAGACGACCTCTCATG	780
QY	781	TACATCCCGCGAGAGGCCAGCGCGCGGACCCAGGCGCAAGAGACAGAGCGCTGCCTCATC	840
Db	781	TACATCCCGCGAGAGGCCAGCGCGCGGACCCAGGCGCAAGAGACAGAGCGCTGCCTCATC	840
QY	841	AGCTAG 846	
Db	841	AGCTAG 846	
RESULT 6	AF177335	1746 bp	linear
LOCUS	AF177335	1746 bp	linear
DEFINITION	Homo sapiens clone SP1942 unknown mRNA.		
ACCESSION	AF177335		
VERSION	AF177335.1	GI:10503968	
KEYWORDS			
SOURCE			
ORGANISM	Homo sapiens.		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	1 (bases 1 to 1746)		
AUTHORS	Gu,J.R., Wan,D.F., Zhao,X.T., Zhou,X.M., Jiang,H.Q., Zhang,P.P.,		
	Qiu,W.X., Huang,Y., Qiu,X.K., Qian,L.F., He,L.P., Li,H.N., Yu,Y.,		
	Yu,J. and Han,L.H.		
TITLE	Novel human cDNA clone with function of inhibiting cancer cell growth		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 1746)		
AUTHORS	Gu,J.R., Wan,D.F., Zhao,X.T., Zhou,X.M., Jiang,H.Q., Zhang,P.P.,		
	Qiu,W.X., Huang,Y., Qiu,X.K., Qian,L.F., He,L.P., Li,H.N., Yu,Y.,		
	Yu,J. and Han,L.H.		
TITLE	Direct Submission		
JOURNAL	Submitted (06-AUG-1999) National Laboratory For Oncogenes & Related		
	Genes, Shanghai Cancer Institute, 25/Ln 2200 Xie Tu Road, Shanghai		
	200032, P.R. China		
FEATURES	location/Qualifiers		
source	1..1746		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="SP1942"		

CDS

455..1039
 /codon_start=-1
 /product="unknown"
 /protein_id="AA017979.1"
 /db_xref="GI:10503969"
 /translation="MRRSLITGDFVILVSLNDRSFEVQRLROOILDTKSLNKK
 TKENVDPVLYICNGKDRDFRYEVDQREIQLVGDDEFORCAITEISAKKNSLDQMR
 ALFEMAKLPSEMSPDLRKRVQYCDVLRKALRNKLIRAGSGGGGGGDFGIVA
 PFRARPSVHSDLMYIRKASAGSQAOKRECVIS"

BASE COUNT 388 a 569 c 496 g 293 t

ORIGIN

Query Match 96.7%; Score 835; DB 9; Length 1746;
 Best Local Similarity 99.9%; Pred. No. 2.3e-102;
 Matches 846; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

1 ATGAAGTGGCGGATGATCAAGAGATGAGCCGAGCGATCGAGTGAATCCCG 60
 193 ATGAAGTGGCGGATGATCAAGAGATGAGCCGAGCGATCGAGTGAATCCCG 252
 61 GCCAAGACTGCTATCGATGATGATCCTCGCTCGTCCAGAGTGGGACAGCGCATC 120
 253 GCCAAGACTGCTATCGATGATGATCCTCGCTCGTCCAGAGTGGGACAGCGCATC 312
 121 GTGTGGCGCTTCTCTACCGCGCGCTTGAGAGCGCTACACGCTACCATGAGACTTC 180
 313 GTGTGGCGCTTCTCTACCGCGCGCTTGAGAGCGCTACACGCTACCATGAGACTTC 372
 181 CACCGCAAGTCTTACTCCATCCGCGCGAGGCTACACAGC-TGCAGATCCCTCGACAGCTC 239
 373 CACCGCAAGTCTTACTCCATCCGCGCGAGGCTACACAGC-TGCAGATCCCTCGACAGCTC 432
 240 CGGCAACACCGCTTCCCGCGCGAGCGCGCTCTCCATCCCTCACAGAGAGCTTTTCAT 299
 433 CGGCAACACCGCTTCCCGCGCGAGCGCGCTCTCCATCCCTCACAGAGAGCTTTTCAT 492
 300 CCTGTGTTCAGTCTGAGCAACCGGATCTCTTGAGAGAGTGCAGGCGCTCAGGACGA 359
 493 CCTGTGTTCAGTCTGAGCAACCGGATCTCTTGAGAGAGTGCAGGCGCTCAGGACGA 552
 360 GATCCTGACACCAAGTCTTCTCAAGCAAAACCAAGAGAGAGTGCAGGCGCTCAGGACGA 419
 553 GATCCTGACACCAAGTCTTCTCAAGCAAAACCAAGAGAGAGTGCAGGCGCTCAGGACGA 612
 420 GGTCTATGCGGCAACAGAGGTGACCGCGACTTCTACCGCGAGGTGACACGCGGAGAT 479
 613 GGTCTATGCGGCAACAGAGGTGACCGCGACTTCTACCGCGAGGTGACACGCGGAGAT 672
 480 CGAGAGCTGTGGGCGGACGACCGCCAGCGCTGGCGCTTCAAGATCTGGGCAAGAA 539
 673 CGAGAGCTGTGGGCGGACGACCGCCAGCGCTGGCGCTTCAAGATCTGGGCAAGAA 732
 540 GAACAGCAGCTGACAGATGTTCCGCGCGCTCTTGCCATGAGCAAGTGCAGGCGGAGGA 599
 733 GAACAGCAGCTGACAGATGTTCCGCGCGCTCTTGCCATGAGCAAGTGCAGGCGGAGGA 792
 600. GATGAGCCAGACCTGACACCGCAAGGCTCTGCTGCAAGTGCAGGCTGCAAGAA 659
 793 GATGAGCCAGACCTGACACCGCAAGGCTCTGCTGCAAGTGCAGGCTGCAAGAA 852
 660 GGGCGTGGGGAACAAGAGTGTGCGGGCGGCAAGGCGCGCGCGCGGCAAGCGGG 719
 853 GGGCGTGGGGAACAAGAGTGTGCGGGCGGCAAGGCGCGCGCGCGGCAAGCGGG 912
 720 CGAGCGCTTGGCATGCGTGGGCAAGCTTGGCGCGCGCGCGCGGCGGCAAGCGCTCAT 779
 913 CGAGCGCTTGGCATGCGTGGGCAAGCTTGGCGCGCGCGCGCGGCGGCAAGCGCTCAT 972
 780 GTATATCGCGAGAGAGCGCGCGCGGCGGCAAGGCGGCAAGGAGAGCGCGCTCGTAT 839
 973 GTATATCGCGAGAGAGCGCGCGCGGCGGCAAGGCGGCAAGGAGAGCGCGCTCGTAT 1032
 840 CAGCTAG 846

Db 1033 CAGCTAG 1039

|||||||

RESULT 7
 BC034166
 LOCUS
 DEFINITION
 ACCESSION
 BC034166
 KEYWORDS
 SOURCE
 ORGANISM

BC034166 1612 bp mRNA linear ROD 07-AUG-2002
 Mus musculus, RAS, dexamethasone-induced 1, clone MGC:36188
 IMAGE:4989312, mRNA, complete cds.
 BC034166 GI:21706874
 MGC.
 house mouse.
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 1612)
 Strausberg, R.
 Direct Submission
 Submitted (02-JUL-2002) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA

REFERENCE
 AUTHORS
 TITLE
 JOURNAL

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabbs-r@mail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome
 Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www-shgc.stanford.edu>
 Contact: (Dickson, Mark) mdedpaxil@stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
 R. M.

REMARK
 COMMENT

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRRAK Plate: 58 Row: 9 Column: 5
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 6677672.
 Location/Qualifiers
 1..1612
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /map="FVB/N"
 /clone="MGC:36188 IMAGE:4989312"
 /tissue_type="Colon, normal, 5 month old male mouse."
 /clone_lib="NCI-CGAP_C024"
 /lab_host="DH10B"
 /note="Vector: PCMV-SF0R6"
 108..950
 /codon_start=-1
 /product="RAS, dexamethasone-induced 1"
 /protein_id="AAH34166.1"
 /db_xref="GI:21706874"
 /db_xref="LocustID:19416"
 /translation="MKLAMIRKMGPSSELSIPKNCRYVILGSSKVGTAIVSR
 LIGREDAVTPRIEDHNRKFSIRGEVQIOLIDISGHPHPPARIRSLILNGPFIIV
 PSLNDRSEEVQRLKOIILDTKSLKRTKENVDPVLYICNGKDRDFRYEVDREI
 EQLVGDDEFORCAITEISAKKNSLDQMRALFEMAKLPSEMSPDLRKRVQYCDVLR
 KALRNKLIRAGSGGGGDFGIVAPFRARPSVHSDLMYIRKTSVGSQAOKDER
 CVIS"

CDS

BASE COUNT 417 a 424 c 420 g 351 t

ORIGIN

Query Match 80.1%; Score 678; DB 10; Length 1612;
 Best Local Similarity 88.4%; Pred. No. 2.3e-81;
 Matches 748; Conservative 0; Mismatches 95; Indels 3; Gaps 1;

1 ATGAAGTGGCGGATGATCAAGAGATGAGCCGAGCGATCGAGTGAATCCCG 60
 108 ATGAAGTGGCGGATGATCAAGAGATGAGCCGAGCGATCGAGTGAATCCCG 167

```

QY 61 GCCAAGACGCTGATTCGCAATGCTATCTCTGCTCTGCTCAAGTGGGCAAGACGCCATC 120
    |||||
Db 168 GCCAAGACGCTGATTCGCAATGCTATCTCTGCTCTGCTCAAGTGGGCAAGACGCCATC 227
QY 121 GTGCGCGCTTCCGACGCGGCTTGGAGAGCGCTACAGCGCTCCACGAGGAGCTTC 180
    |||||
Db 228 GTGCGCGCTTCCGACGCGGCTTGGAGAGCGCTTACAGCGCTCCACGAGGAGCTTC 287
QY 181 CACCGCAAGTTCTACTCATCCGCGGAGGTGTACAGCTGCAATCTCGACAGCTC 240
    |||||
Db 288 CACCGCAAGTTCTACTCATCCGCGGAGGTGTACAGCTGCAATCTCGACAGCTC 347
QY 241 GGCAGACACCCGTTCCCGGCAATGCGGCGCTCTCATCTCTCACAGAGACGTTTTCATC 300
    |||||
Db 348 GGCAGACACCCGTTCCCGGCAATGCGGCGCTCTCATCTCTCACAGAGACGTTTTCATC 407
QY 301 CTGATGTTCTGCTGCAACCGGCACTCTTGGAGAGGTGTGAGGCGCTGACGAGCGAG 360
    |||||
Db 408 CTGATGTTCTGCTGCAACCGGCACTCTTGGAGAGGTGTGAGGCGCTGACGAGCGAG 467
QY 361 ATCTCTGACACCAAGTCTGCTCTCAAGAACCAAGAGAGAACTGGAAGTGCCTC 420
    |||||
Db 468 ATCTCTGACACCAAGTCTGCTCTCAAGAACCAAGAGAGAACTGGAAGTGCCTC 527
QY 421 GTGATCTGCGGCAACCAAGGTGTACCGGAGTCTTACCGGAGGTGTGACAGCGCGAGATC 480
    |||||
Db 528 GTGATCTGCGGCAACCAAGGTGTACCGGAGTCTTACCGGAGGTGTGACAGCGCGAGATC 587
QY 481 GAGCAGCTGTGTGGGCGACGACCGCGCTGCGCTGCTGAGATCTGCGGCAAGAG 540
    |||||
Db 588 GAGCAGCTGTGTGGGCGACGACCGCGCTGCGCTGCTGAGATCTGCGGCAAGAG 647
QY 541 AACAGCAGCTGACCAAGATGTTCCGCGGCTCTTGCAGTGGCAAGTGCCTCAGCGAG 600
    |||||
Db 648 AACAGCAGCTGACCAAGATGTTCCGCGGCTCTTGCAGTGGCAAGTGCCTCAGCGAG 707
QY 601 ATGAGCGCAACCGTGCAGCGGCAAGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
    |||||
Db 708 ATGAGCGCAACCGTGCAGCGGCAAGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 767
QY 661 GCGCTGTGGGCAACCAAGGTGTGTGCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 720
    |||||
Db 768 GCGCTGTGGGCAACCAAGGTGTGTGCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 824
QY 721 GAGCGCTTGGCAATCTGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 780
    |||||
Db 825 GAGCGCTTGGCAATCTGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 884
QY 781 TACATCGCGGAGAAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 840
    |||||
Db 885 TACATCGGTAAGAAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 944
QY 841 AGCTAG 846
    |||||
Db 945 AGTTAG 950

```

```

RESULT 8
AF239157 1616 bp mRNA linear ROD 06-JUL-2001
LOCUS AF239157
DEFINITION Rattus norvegicus DEXRAS1 (Dextral) mRNA, complete cds.
ACCESSION AF239157
VERSION AF239157.1 GI:7230767
KEYWORDS
SOURCE Rattus norvegicus.
ORGANISM Rattus norvegicus.
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
          Rattus.
REFERENCE 1 (bases 1 to 1616)
AUTHORS Fang, M., Jaffrey, S.R., Sawa, A., Ye, K., Luo, X. and Snyder, S.H.
TITLE Dextral: a G protein specifically coupled to neuronal nitric oxide

```

```

JOURNAL synthase via CAPON
MEDLINE Neuron 28 (1), 183-193 (2000)
PUBMED 20537828
PUBMED 11086993
REFERENCE 2 (bases 1 to 1616)
AUTHORS Fang, M., Jaffrey, S.R., Sawa, A., Ye, K. and Snyder, S.H.
TITLE Direct Submission
JOURNAL Submitted (24-FEB-2000) Neuroscience, Johns Hopkins University, 725
          N. Wolfe St, Baltimore, MD 21205, USA
FEATURES
source 1..1616
          /organism="Rattus norvegicus"
          /db_xref="taxon:10116"
          /gene="Dextral"
          /note="has-related protein; small GTPase; similar to Mus
          musculus DEXRAS1 encoded by GenBank accession Number
          AF09246"
          /product="Dextral"
          /protein_id="AAFA3090.1"
          /db_xref="GI:7230768"
          /translation="MKLAAMTKKCPDSDELSPKNCYRMVILIGSKVGKTAIVSF
          LTREREDAVPTLIDFPRKFSIRGEVYQDIDTSONHPFPAARRLSILTGVEFLV
          FSLNDRDSFEVORIKQOIIDTKSLKNEVDPLVIGKNGDRDFREYREI
          EQLVDDPDRCAVEISAKRSISLDORFALFMAKLPSKSPDLHRKVSVOYCDVH
          KKALRNKLLRAGSGGGGHDGADFGIAPFARRRSVSHDLITREKTSVSQAKDKER
          CVIS"
BASE COUNT 403 a 436 c 413 g 363 t 1 others
ORIGIN
Query Match 80.1%; Score 678; DB 10; Length 1616;
Best Local Similarity 88.4%; Pred. No. 2,3e-81;
Matches 748; Conservative 0; Mismatches 95; Indels 3; Gaps 1;
QY 1 ATGAAGCTGGCGGCGGCAAGATGTTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 60
    |||||
Db 118 ATGAAGCTGGCGGCGGCGGCAAGATGTTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 177
QY 61 GCGAAGCTGTATCGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
    |||||
Db 178 GCGAAGCTGTATCGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 237
QY 121 GTGCGCGCTTCCGACGCGGCTTGGAGAGCGCTACAGCGCTCCACGAGGAGCTTC 180
    |||||
Db 238 GTGCGCGCTTCCGACGCGGCTTGGAGAGCGCTTACAGCGCTCCACGAGGAGCTTC 297
QY 181 CACCGCAAGTTCTACTCATCCGCGGAGGTGTACAGCTGCAATCTCGACAGCTTC 240
    |||||
Db 298 CACCGCAAGTTCTACTCATCCGCGGAGGTGTACAGCTGCAATCTCGACAGCTTC 357
QY 241 GGCAGACACCCGTTCCCGGCAATGCGGCGCTCTCATCTCTCACAGAGACGTTTTCATC 300
    |||||
Db 358 GGCAGACACCCGTTCCCGGCAATGCGGCGCTCTCATCTCTCACAGAGACGTTTTCATC 417
QY 301 CTGATGTTCTGCTGCAACCGGCACTCTTGGAGAGGTGTGAGGCGCTGACGAGCGAG 360
    |||||
Db 418 CTGATGTTCTGCTGCAACCGGCACTCTTGGAGAGGTGTGAGGCGCTGACGAGCGAG 477
QY 361 ATCTCTGACACCAAGTCTGCTCTCAAGAACCAAGAGAGAACTGGAAGTGCCTC 420
    |||||
Db 478 ATCTCTGACACCAAGTCTGCTCTCAAGAACCAAGAGAGAACTGGAAGTGCCTC 537
QY 421 GTGATGTTCTGCTGCAACCGGCACTCTTGGAGAGGTGTGAGGCGCTGACGAGCGAG 480
    |||||
Db 538 GTGATGTTCTGCTGCAACCGGCACTCTTGGAGAGGTGTGAGGCGCTGACGAGCGAG 597
QY 481 GAGCAGCTGTGTGGGCGACGACCGCGCTGCGCTGCTGAGATCTCGGCAAGAG 540
    |||||
Db 598 GAGCAGCTGTGTGGGCGACGACCGCGCTGCGCTGCTGAGATCTCGGCAAGAG 657

```

Query	Match	80.0%	Score	676.4	DB	10	Length	1623
541	AACACAGCCTTGAGCACAATGTTTCGCGCGCTCTTCGCGCATGCGCAAGTCCACGCGAG	600						
658	AATACAGCCTTGAGCACAATGTTTCGCGCGCTCTTCGCGCATGCGCAAGTCCACGCGAG	717						
601	ATGACCCCAAGACCTGCACCGCAAGGTTCTCGGTGACAGTACTGCGACGTGCTGCACAAGAG	660						
718	ATGACCCCTTGACCTTCACCGCAAGGTTGCTGTGAGTACTGTGACGTCTGTCACAAAAAG	777						
661	GGCTGCGGGAACAAAGAAGTGTGCGGGCGGCGACGGGCGGCGGGCGGCGACCCGGGAG	720						
778	GCTCTGAGGAAACAAAGAAGTGTGCGGGCGGCGACGGGCGGCGGGCGGCGACCCGGGAG	834						
721	GACGCGCTTTGGCATTGCGGACACCTTCGCGCGCGCGGCGGCGGCGGACGACGACCTCATG	780						
835	GATGCTTTGGCATCTTGCGCGCCCTTTGCTGCTGCAGAACCTAAGCGTCATAGCAGCCTCATG	894						
761	TACATCCCGGAGAAAGCCGACGCGCGGCGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	840						
895	TACATCTCGTGAAGAAACCAAGTGTGACAGCCAGCGCTAAGGACAAAGGACGCGTGTGTATC	954						
841	AGCTAG	846						
955	AGTAG	960						
RESULT 9								
AF009246	1623 bp	mRNA	linear	ROD 13-FEB-1998				
LOCUS	Mus musculus ras-related protein (DEXRAS1) mRNA, complete cds.							
DEFINITION	Mus musculus ras-related protein (DEXRAS1) mRNA, complete cds.							
ACCESSION	AF009246							
VERSION	AF009246.1	GI:2253712						
KEYWORDS								
SOURCE	Mus musculus.							
ORGANISM	Mus musculus							
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;							
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.							
REFERENCE	1 (bases 1 to 1623)							
AUTHORS	Kempainen,R.J. and Behrend,E.N.							
TITLE	Dexamethasone rapidly induces a novel ras superfamily							
	member-related gene in AtT-20 cells							
JOURNAL	J. Biol. Chem. 273 (6), 3129-3131 (1998)							
MEDLINE	98123070							
PUBMED	9452419							
REFERENCE	2 (bases 1 to 1623)							
AUTHORS	Kempainen,R.J. and Behrend,E.N.							
TITLE	Direct Submission							
JOURNAL	Submitted (18-JUN-1997) Physiology & Pharmacology, Auburn							
	University College of Veterinary Medicine, 213 Greene Hall, Auburn,							
	AL 36849, USA							
FEATURES								
source	location/Qualifiers							
	1..1623							
	/organism="Mus musculus"							
	/db_xref="taxon:10090"							
	/cell_line="corticotrope tumor cell line AtT-20"							
	1..1623							
	/gene="DEXRAS1"							
	142..984							
	/gene="DEXRAS1"							
	/note="Induced by dexamethasone"							
	/codon_start=1							
	/product="ras-related protein"							
	/protein_id="AAC53538.1"							
	/db_xref="GI:2253713"							
	/translation="MKLAAMTKKMCPSDELSTIPKNCYRMVITIGSSVNGKTAIVSEF							
	LTGREDEATPTTIEDFHKFTKSTIGEYVLDIDLTSGNHPPARRKSLTGDVFIIV							
	FSLDNRDSFEVORLKOQILDFKSLCKNKTENVDVPLTGNGGDHDFYEVSRQRELT							
	EQVLDGDDPQRCAYFEISAKKNSLDQMRALFMAAKLPSEMSPLHRRVSVQYDVVHL							
	KKALRNKKLLRAGSGGGGDDHDAFGIILAFARRPSVSDMLYIREKTSVGSQADKDEE							

	Best Local Similarity	88.3%	Pred. No. 3,76-81:	
	Matches	747:	Conservative	0; Mismatches 96; Indels 3; Gaps 1
QY	1	ATGAACTGGCCCGCATGATCAAGAAAGATGTCCCGGACCACTCGGAGCTGATATCCG	60	
Db	142	ATGAAACTGGCCCGCATGATCAAGAAAGATGTCCCGGACCACTCGGAGCTGATATCCG	201	
QY	61	GCCAAAGATGCTATGCGATGTCATCTCGGCTGTCCCAAGGTGGGCAAGAGCCATC	120	
Db	202	GCCAAAGATGCTATGCGATGTCATCTCGGCTGTCCCAAGGTGGGCAAGAGCCATC	261	
QY	121	GTGTGCGCTCTTCTCAACCGGCGCTTTCGAGAGCGCTTACACGCTTACATGAGGACTTC	180	
Db	262	GTGTGCGCTCTTCTCAACCGGCGCTTTCGAGAGCGCTTACACGCTTACATGAGGACTTC	321	
QY	181	CACCGCAAGTTCTACTCCATCCGCGGCGAGGTCCTACAGCTCGACATCTCTGACAGCTCC	240	
Db	322	CACCGCAAGTTTACTCTGATCCGCGGCGAGGTCCTACAGCTCGACATCTGACAGCTCC	381	
QY	241	GGCAGACACCCGCTTCCCGCATGCGGCGCTCTCTCATCTCTCACAGAGAGAGCTTTTCATC	300	
Db	382	GGCAATCATCCGTTTCCCGCATGCGGCGCTCTCTCATCTCTCACAGAGAGAGTTTTCATC	441	
QY	301	CTGCTGTTCACTGTGAGACAACCGCACTCTTTCGAGAGGCTGACGCGCTTCAGCAGAC	360	
Db	442	CTGCTGTTCACTGTGAGACAACCGCACTCTTTCGAGAGGCTGACGCGCTTCAGCAACAC	501	
QY	361	ATTCCTCACACCAAGTCTTGCCCTCAAGAACCAAGAGAGAACGTTGAGAGCTGCCCTG	420	
Db	502	ATTCCTCACACCAAGTCTTGCTCTCAAGAACCAAGAGAGAGATGTGAGAGTGCCTG	561	
QY	421	GTCATCTGGCGCAACAGGGTGAACGCGCACTTCTACCGCGAGGTTGACCAAGCGGAGATC	480	
Db	562	GTCATTTGGCGTAAACAGGGGACCGGCACTTCTACCGGAGAGTGAAGACGCGGAGATT	621	
QY	481	GAGCAGTGGTGGGGCGACGACCCCGAGCGCTGCGCTTCTCGAGATCTGTGGCCAAAGAG	540	
Db	622	GAGCAGCTTGGTGGTGAAGCCCTTCAAGCTTGTGCTTCTTCTGAGATCTCAGCCAGAGAG	681	
QY	541	AACAGCAGCTTGGACCAAGATGTTTCGCGGCGCTTTCGGCATGAGCCAGCTGCCAGGAG	600	
Db	682	AACAGCAGCTTGGACCAAGATGTTTCGCTTGGCATTGGCCATGAGCCAGCTGCCAGGAG	741	
QY	601	ATGAGCCCAACCTGCAACCGCAAGGTCTGTGGTGCAGTACTGCGACGTGCTGCACAGAG	660	
Db	742	ATGAGCCCAACCTGCAACCGCAAGGTATGTGCAGTACTGCGACGTGCTGCACAGAGAG	801	
QY	661	GCGCTGGGGAACAAGAGCTGTGCGGGGCGGCGACGCGCGCGCGCGCGAGCCGGG	720	
Db	802	GCTGTGAGGAACAAGAGCTGTGCGGGGCGGCGACGCGCGCGCGCGCGAGCCGGG	858	
QY	721	GAGCGCTTGTGGCATCTGTGCAACCTTTCGCGCGCGCGCGCGCGCTGTACACAGGACTTCAG	780	
Db	859	GATCCCTTTTGGCATCTTGTGGCGCGCTTTCGCTCCCAACACCGCTGACAGGACTTCAG	918	
QY	781	TACATTCGCGAGAGGCGACGCGCGCGCACCGAGGCGCAAGAGAGAGCGCTGTGCTATC	840	
Db	919	TACATTTGTAAGAAAAACCAAGTGTGCGACCGAGGCTTAAGCAAGAGAGCGCTGTGCTATC	978	
QY	841	AGCTAG 846		
Db	979	AGTTAG 984		
RESULT 10				
LOCUS	AF222979	4990 bp	DNA	linear PRI 02-JAN-2001
DEFINITION	Homo sapiens activator of G-protein signaling gene, complete cds.			
ACCESSION	AF222979			
VERSION	AF222979.1	GI:12004991		
KEYWORDS				
SOURCE	Homo sapiens			
ORGANISM	Homo sapiens			

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE
 1 (bases 1 to 4990)
 AUTHORS
 Cismowski, M.J., Xie, X. and Duzic, E.
 TITLE
 Genomic sequence of the human ras-related G-protein activator AGS1
 JOURNAL
 Unpublished
 REFERENCE
 2 (bases 1 to 4990)
 AUTHORS
 Cismowski, M.J., Xie, X. and Duzic, E.
 TITLE
 Direct Submission
 Submitted (10-JAN-2000) OSI Pharmaceuticals, 777 Old Saw Mill River
 Road, Tarrytown, NY 10591, USA
 JOURNAL
 Location/Qualifiers
 FEATURES
 source
 1. 4990
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="17"
 /map="17p11.2"
 /clone="BamHI-EcoRI subclone of p1"
 /note="chromosomal localization verified by fluorescence
 in-situ hybridization analysis"
 join(2762..3188,3400..4643)
 /product="activator of G-protein signaling"
 join(2903..3188,3400..3959)
 /note="ras-related protein; AGS1"
 /codon_start=1
 /product="activator of G-protein signaling"
 /protein_id="AA644256.1"
 /db_xref="GI:12004992"
 /translation="MTAMIKKCPSELSIPANKRYVILSSKVGKTAIYSRF
 LGRFEDATPTIEDPHRFYSIRGEVOLDILIDNSPPFAMKSLITGDVFLV
 FSLDNDSFEEOYRLROOILDPKSCILKNTKENVVPLVYCNKNDROPFRVDDREI
 EOLVDDPDRCAYPELSAKNSLSDOMRALPMKLESESPDLHKKVSVYCVLH
 KVALRKKLRLRAGSGGGGDPDAFGLVAFARPSVSHDLVIREKASAGSQARDE
 RCVIS"
 BASE COUNT 1049 a 1507 c 1509 g 925 t
 ORIGIN
 Query Match 73.9%; Score 625; DB 9; Length 4990;
 Best Local Similarity 80.0%; Pred. No. 2.1e-74;
 Matches 846; Conservative 0; Mismatches 0; Indels 211; Gaps 1;

QY 285 -----AGCAGACGTTTTCATCTGCTGTTCACTGCGAACACCGGACTC 329
 Db 3383 CTCTGTGCCCCCTCTAGAGAGAGCTTTTCATCTGCTGTTCACTGCGAACACCGGACTC 3442
 QY 330 CTTCGAGAGGTGACGCGGCTCAGGACAGATCTCTCACACCAAGTTGCTTCAGAA 389
 Db 3443 CTTCGAGAGGTGACGCGGCTCAGGACAGATCTCTCACACCAAGTTGCTTCAGAA 3502
 QY 390 CAAACCAAGGAAAGTGGACGTGCCCCCTGTCATCTGCGCAACAGGGTGAACCGGA 449
 Db 3503 CAAACCAAGGAAAGTGGACGTGCCCCCTGTCATCTGCGCAACAGGGTGAACCGGA 3562
 QY 450 CTTCACGCGGAGGTGGACACGCGAATGACAGCTGTGTGGCGACACCCGACG 509
 Db 3563 CTTCACGCGGAGGTGGACACGCGAATGACAGCTGTGTGGCGACACCCGACG 3622
 QY 510 CTGCGCTACTCTGAGATCTGCGCCCAAGAAACGACCTGGACGATGTTCCGCGC 569
 Db 3623 CTGCGCTACTCTGAGATCTGCGCCCAAGAAACGACCTGGACGATGTTCCGCGC 3682
 QY 570 GCTCTTCCGATGGCCAAAGTGGCCGAGATGAGCCAGACCTGCAACCGAGGTCTC 629
 Db 3683 GCTCTTCCGATGGCCAAAGTGGCCGAGATGAGCCAGACCTGCAACCGAGGTCTC 3742
 QY 630 GTTGCACTACTGCGACGTGCTGACACAGAGCGCTGCGGAACAAGCTGTGCGGC 689
 Db 3743 GTTGCACTACTGCGACGTGCTGACACAGAGCGCTGCGGAACAAGCTGTGCGGC 3802
 QY 690 CGCAGCG 749
 Db 3803 CGCAGCG 3862
 QY 750 GCG 809
 Db 3863 GCG 3922
 QY 810 CCAGGCCAAGGCAAGGAGCGCTGCTCATCAGCTAG 846
 Db 3923 CCAGGCCAAGGCAAGGAGCGCTGCTCATCAGCTAG 3959
 RESULT 11
 AC020558/c 183334 bp DNA linear PRI 09-JAN-2002
 LOCUS
 DEFINITION Homo sapiens BAC clone Rpl1-524F11 from 17, complete sequence.
 AC020558
 VERSION
 KEYWORDS
 HTG.
 SOURCE
 ORGANISM
 Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE
 1 (bases 1 to 183334)
 AUTHORS
 Sulston, J.E. and Waterston, R.
 TITLE
 Toward a complete human genome sequence
 JOURNAL
 Genome Res. 8 (11), 1097-1108 (1998)
 MEDLINE
 PUBMED
 9847074
 REFERENCE
 2 (bases 1 to 183334)
 AUTHORS
 Mulvaney, E., Maupin, R., Laplant, Y. and Bielicki, L.
 TITLE
 The sequence of Homo sapiens BAC clone Rpl1-524F11
 JOURNAL
 Unpublished (2001)
 REFERENCE
 3 (bases 1 to 183334)
 AUTHORS
 Waterston, R.H.
 TITLE
 Direct Submission
 Submitted (03-JAN-2000) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 REFERENCE
 4 (bases 1 to 183334)
 AUTHORS
 Waterston, R.H.
 TITLE
 Direct Submission
 Submitted (07-MAR-2001) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,

REFERENCE MO 63108, USA
 5 (bases 1 to 183334)
 AUTHORS Waterston, R.
 TITLE Direct Submission
 JOURNAL Submitted (09-JAN-2002) Department of Genetics, Washington University, 444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 On Mar 7, 2001 this sequence version replaced g1:7631041.

----- Genome Center
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: <http://genome.wustl.edu/gsc>
 Contact: sapiens@wustl.wustl.edu
 ----- Summary Statistics
 Center project name: H_NH0524F11

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
 Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:
 The RPCR-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P.Y., Zhao, B., Frengen, E., Tatematsu, M., Catanese, J.J., and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://dpcpec.med.buffalo.edu>)
 VECTOR: pBAC3.6
 NEIGHBORING SEQUENCE INFORMATION:
 The clone sequenced to the left is AC073621. Actual start of this clone is at base position 1 of RPl1-524F11; actual end is at base position 183334 of RPl1-524F11.

The sequence from position 1987 to 2253 was derived from PCR product of RPl1-524F11 BAC DNA.

FEATURES
 source
 Location/Qualifiers
 1..183334
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="17"
 /map="17"
 /clone="RPl1-524F11"
 /clone_id="RPCR-11"
 1..319

misc_feature
 /note="similar to EST A1672824 (NID:g4852555) we73e06.x1"
 339..887
 misc_feature
 /note="similar to EST A1968384 (NID:g5765202) wu02e05.x1"
 339..738
 misc_feature
 /note="similar to EST A1289781 (NID:g39331622) gw12a07.x1"
 339..658
 misc_feature
 /note="similar to EST AW072137 (NID:g6027135) w298e07.x1"
 340..767
 misc_feature
 /note="similar to EST A1014398 (NID:g3228779) ou93b12.s1"
 340..686
 misc_feature
 /note="similar to EST AW237782 (NID:g6570171) xm81e08.x1"
 346..742

misc_feature
 /note="similar to EST A1359064 (NID:g4110685) qy25e07.x1"
 405..884
 misc_feature
 /note="similar to EST A1279211 (NID:g3917445) qm24c09.x1"
 782..1335
 misc_feature
 /note="similar to EST BE042464 (NID:g10759519)"
 822..1319
 misc_feature
 /note="similar to EST BE864687 (NID:g10385980)"
 834..2931
 misc_feature
 /note="CPG island (%GC=69.0, o/e=0.80, #CPGs=226)"
 1135..1585
 misc_feature
 /note="similar to EST BE555582 (NID:g11665312)"
 1168..1583
 misc_feature
 /note="similar to EST AW028127 (NID:g5886883) wv26c07.x1"
 1432..1583
 misc_feature
 /note="similar to EST BE390509 (NID:g9335874)"
 1440..1583
 misc_feature
 /note="similar to EST AW915326 (NID:g8081017)"
 1450..1583
 misc_feature
 /note="similar to EST AA790463 (NID:g2850583) vw04g12.r1"
 1795..2255
 misc_feature
 /note="similar to EST BE390509 (NID:g9335874)"
 1795..2201
 misc_feature
 /note="similar to EST AA790463 (NID:g2850583) vw04g12.r1"
 1795..2153
 misc_feature
 /note="similar to EST AW915326 (NID:g8081017)"
 1795..2114
 misc_feature
 /note="similar to EST AW028127 (NID:g5886883) wv26c07.x1"
 1804..2071
 misc_feature
 /note="similar to EST BE836084 (NID:g10268462)"
 1817..2201
 misc_feature
 /note="similar to EST AA672515 (NID:g2644732) vo59e02.r1"
 1844..2201
 misc_feature
 /note="similar to EST AW463433 (NID:g7033601)"
 1985..2285
 misc_feature
 /note="similar to EST AA852097 (NID:g2940690)"
 2538..3033
 misc_feature
 /note="similar to EST BF057610 (NID:g10811506)"
 2554..2806
 misc_feature
 /note="similar to EST BF195653 (NID:g11082770)"
 3561..3924
 misc_feature
 /note="similar to EST R18261 (NID:g771871) yg01a05.r1"
 4027..4085
 misc_feature
 /rpl_family="MERL_type"
 4670..5117
 misc_feature
 /note="similar to EST R42241 (NID:g817108) yf98b09.s1"
 4857..5117
 misc_feature
 /note="similar to EST T16275 (NID:g518437)"
 6155..6307
 repeat_region
 /rpl_family="Alu"
 6309..6446
 repeat_region
 /rpl_family="L2"
 6583..6891
 repeat_region
 /rpl_family="Alu"
 6960..7153
 repeat_region
 /rpl_family="L1"
 7525..7838
 repeat_region
 /rpl_family="Alu"
 7851..8487
 repeat_region
 /rpl_family="L1"
 9000..9104
 repeat_region
 /rpl_family="L2"
 11463..11750
 misc_feature
 /note="similar to EST BF344873 (NID:g11292331)"
 11465..11750
 misc_feature
 /note="similar to EST A1340185 (NID:g4077112) qo79d07.x1"
 11465..11750
 misc_feature
 /note="similar to EST A1571390 (NID:g4534764) tn45e12.x1"
 11465..11750
 misc_feature
 /note="similar to EST A1701138 (NID:g4989038) we10d09.x1"
 11465..11750
 misc_feature
 /note="similar to EST A1803637 (NID:g5369109) tc43d01.x1"
 11465..11750
 misc_feature
 /note="similar to EST BE747895 (NID:g10161887)"

```

misc_feature      11465..11750
                   /note="similar to EST BF203429 (NID:g1097015)"
misc_feature      11465..11697
                   /note="similar to EST A1869572 (NID:g5543540) w198d03.x1"
misc_feature      11467..11750
                   /note="similar to EST A1817466 (NID:g5436545) wk38g01.x1"
misc_feature      11467..11619
                   /note="similar to EST AA367144 (NID:g2019483)"
misc_feature      11468..11750
                   /note="similar to EST BE218333 (NID:g8905651) hv37c09.x1"
misc_feature      11468..11750
                   /note="similar to EST BE549542 (NID:g9791256)"
misc_feature      11468..11750
                   /note="similar to EST BE750272 (NID:g10164264)"
misc_feature      11478..11750
                   /note="similar to EST BF667874 (NID:g11941769)"
misc_feature      11480..11717
                   /note="similar to EST R34866 (NID:g791767) yh86f06.s1"
misc_feature      11483..11750
                   /note="similar to EST AA429062 (NID:g2110588) zv49a05.r1"
misc_feature      11541..11750
                   /note="similar to EST W38322 (NID:g1320126) zb19g06.r1"
misc_feature      11552..11750
                   /note="similar to EST BE910310 (NID:g10406772)"

Query Match      73.9%; Score 625; DB 9; Length 183334;
Best Local Similarity 80.0%; Pred. No. 1e-74; 0; Indels 211; Gaps 1;
Matches 846; Conservative 0; Mismatches 0;

QY      1  ATCAAACTGGCGCGATGATCAAGAAATGTGCCCCGAGCGAGCTGAGTATCCCG 60
        |||||||
Db      2080 ATCAAACTGGCGCGATGATCAAGAAATGTGCCCCGAGCGAGCTGAGTATCCCG 2021

QY      61  GCCAAGAACTGCTATGCTATGCTATCTCTGCTGCTGCTGCTCAAGTGGGAAAGCGGCGATC 120
        |||||||
Db      2020 GCCAAGAACTGCTATGCTATGCTATCTCTGCTGCTGCTCAAGTGGGAAAGCGGCGATC 1961

QY      121  GTGTGGCGCTTCCCAACGGCGGCTTCCGAGGAGCGCTACACGCCCTACATCGAGGAGATTC 180
        |||||||
Db      1960 GTGTGGCGCTTCCCAACGGCGGCTTCCGAGGAGCGCTACACGCCCTACATCGAGGAGATTC 1901

QY      181  CACCGCAAGTTCTACTCCATCCGCGGAGAGTCTACACAGCTGACACATCTCTGACACAGTTC 240
        |||||||
Db      1900 CACCGCAAGTTCTACTCCATCCGCGGAGAGTCTACACAGCTGACACATCTCTGACACAGTTC 1841

QY      241  GGCACCAACCGCTTCCCGCGCATGGCGGCTCTCCATCTCTACAGTGAAGCGGCGGCGC 284
        |||||||
Db      1840 GGCACCAACCGCTTCCCGCGCATGGCGGCTCTCCATCTCTACAGTGAAGCGGCGGCGC 1781

QY      285  -----
        |||||||
Db      1780 GGCAGGTGGCGGAGGGAAGGGGGGGAACCTTCGCGCAGGAGCGCCCGGAGCGCGGT 1721

QY      285  -----
        |||||||
Db      1720 CCGGCTGCCGCGCGCGAGTAGTGGCTTGGCGCTTGAAGAGGCTAGCGGCGCCCGCGCG 1661

QY      285  -----
        |||||||
Db      1660 GCCTCAAAAGTCAGCCCACTTGTCCCGGGCGGCGACACCTCACCTTCTCTCTTCTGCT 1601

QY      285  -----
        |||||||
Db      1600 CTCTGTGCCCCCTTAGAGGAGCTTTCATCTGTGTGTCAGTCTGGACAAACCGCGACTC 1541

QY      330  CTTTCGAGAGGTGACAGCGCTCAGCAGACAGATCTCTGACACCAAGTCTTCTCAAGAA 389
        |||||||
Db      1540 CTTTCGAGAGGTGACAGCGCTCAGCAGACAGATCTCTGACACCAAGTCTTCTCAAGAA 1491

QY      390  CAAACCAAGAGAGAGTGAAGTGGTGGTCTCATCTGGGCGACAAAGGGTGAACCGCGA 449
        |||||||
Db      1480 CAAACCAAGAGAGAGTGAAGTGGTGGTCTCATCTGGGCGACAAAGGGTGAACCGCGA 1421

```

```

QY      450  CTTCTACCCGAGGTGAGACCGGCGAGATCGACAGCTGTGTGGCGAGCAGACCCCGAGC 509
        |||||||
Db      1420 CTTCTACCCGAGGTGAGACCGGCGAGATCGACAGCTGTGTGGCGAGCAGACCCCGAGC 1361

QY      510  CTGCGCTTACTTGAAGTCTCGGCGCAAGAAAGACACAGCCTGAGACAGATCTTCGCGC 569
        |||||||
Db      1360 CTGCGCTTACTTGAAGTCTCGGCGCAAGAAAGACACAGCCTGAGACAGATCTTCGCGC 1301

QY      570  GCTCTGCGCATGTGCGCAAGCTGCGCAGCAGATGAGAGCCAGACCTGACCCGAGAGTCTC 629
        |||||||
Db      1300 GCTCTGCGCATGTGCGCAAGCTGCGCAGCAGATGAGAGCCAGACCTGACCCGAGAGTCTC 1241

QY      630  GGTGCACTATGCGACGCTGCTGCGCAAGAAAGCGGCTGCGGAAAGAGTGTGCGCGC 689
        |||||||
Db      1240 GGTGCACTATGCGACGCTGCTGCGCAAGAAAGCGGCTGCGGAAAGAGTGTGCGCGC 1181

QY      690  CGGCAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 749
        |||||||
Db      1180 CGGCAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1121

QY      750  GCGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 809
        |||||||
Db      1120 GCGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1061

QY      810  CCAGGCGAAGAGAGAGGAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 846
        |||||||
Db      1060 CCAGGCGAAGAGAGAGGAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1024

RESULT 12
AP62018      5141 bp      DNA      linear      PRI 31-MAY-2000
LOCUS      AF262018
DEFINITION      Homo sapiens dexamethasone-induced ras-related protein 1 (DEXRAS1)
ACCESSION      AF262018
VERSION      AF262018.1 GI:8118456
KEYWORDS      gene, complete cds.
SOURCE      Homo sapiens.
ORGANISM      Homo sapiens.
REFERENCE      1. (bases 1 to 5141)
AUTHORS      Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE      Human Dexas1 gene
JOURNAL      Unpublished
REFERENCE      2 (bases 1 to 5141)
AUTHORS      Kempainen R.J. and Behrend E.N.
TITLE      Direct Submission
JOURNAL      Submitted (30-Apr-2000) Anatomy & Physiology, Auburn University,
              College of Veterinary Medicine, Auburn, AL 36832, USA

FEATURES
     source
     gene
     mRNA
     CDS

```

	RCVIS"			
BASE COUNT	1090 a	1579 c	1522 g	950 t
ORIGIN				

Query Match	73.7%;	Score 623.4;	DB 9;	Length 5141;
Best Local Similarity	79.9%;	Pred. No. 3.4e-74;		
Matches 845;	Conservative	0;	Mismatches 1;	Indels 211.

QY	1	ATGAAACTGGCCGCGATGATCAAGAAAGTGTGCCGAGCGACTGGAGCTGAGTATCCG	60
Db	2548	ATGAAACTGGCCGCGATGATCAAGAAAGTGTGCCGAGCGACTGGAGTATCCG	60
QY	61	GCCAAAGTGTATTCGATGATGATCTCGGCTCTGTCGAAAGTGGGCAAGAGCGCATC	120
Db	2608	GCCAAAGTGTATTCGATGATGATCTCGGCTCTGTCGAAAGTGGGCAAGAGCGCATC	120
QY	121	GTGTGGCGCTTCTTCACCGGCGCTTGAGAGACCTACAGGCTTACATCGAGACTTC	180
Db	2668	GTGTGGCGCTTCTTCACCGGCGCTTGAGAGACCTACAGGCTTACATCGAGACTTC	180
QY	181	CACGCAAGTTCCTACTCATCCTCGGCGAGGTCTACAGCTGCATCTCGACAGTCC	240
Db	2728	CACGCAAGTTCCTACTCATCCTCGGCGAGGTCTACAGCTGCATCTCGACAGTCC	240
QY	241	GGAACAACCCGTTCCCGCGATGCGCGCGCTCTCATCTCAC-----	284
Db	2788	GGAACAACCCGTTCCCGCGATGCGCGCGCTCTCATCTCACAGGTGAGCCGGGGCC	284
QY	285	-----	284
Db	2848	GGGCAGGTGCGGGAGGAGGAAAGCGGGGAAACCTCGGCGAGGGCGCCCGCAGCCCGCT	2907
QY	285	-----	284
Db	2908	CCGGCTCGCGGGCGCGAGTAGTGCCTTGCCTTAGAGAGCTAGCGCCCGCGCG	2967
QY	285	-----	284
Db	2968	GCCCTAAAGTCAAGCCGAGCTGTCCCTGGGGGCAACCCCTCTCTCTTTCTGCT	3027
QY	285	-----AGGAGAGTTTTCATCTGTGTTCAGTCTGGAACAACCGGACCTC	3027
Db	3028	CTCTGTGCCCCCTTAGAGAGAGTTTTCATCTGTGTTCAGTCTGGAACAACCGGACCTC	3087
QY	330	CTTCGAGAGGTGACAGCGCTCAGGACAGATCCTGACACCAAGCTTGCTCAAGAA	389
Db	3088	CTTCGAGAGGTGACAGCGCTCAGGACAGATCCTGACACCAAGCTTGCTCAAGAA	389
QY	390	CAAAACAAGAGAGAGAGTGGAGAGTCCCTGTATCTCGGGCAACAAGGTGACCGCA	449
Db	3148	CAAAACAAGAGAGAGTGGAGAGTCCCTGTATCTCGGGCAACAAGGTGACCGCA	449
QY	450	CTTCTACCGCGAGGTGGACACAGCGAGATGAGACACTGTGGGGGAGAGACCCGACG	509
Db	3208	CTTCTACCGCGAGGTGGACACAGCGAGATGAGACACTGTGGGGGAGAGACCCGACG	509
QY	510	CTGGGCTACTTCCAGATCTGTGGCAAGAAACAGCAGCTTGAGACATATCTCGCGC	569
Db	3268	CTGGGCTACTTCCAGATCTGTGGCAAGAAACAGCAGCTTGAGACATATCTCGCGC	569
QY	570	GCTCTTCCCATGAGCAAGCTGACCAAGCGAGATGAGACCCAGACCTGACCGCAAGTCTC	629
Db	3328	GCTCTTCCCATGAGCAAGCTGACCAAGCGAGATGAGACCCAGACCTGACCGCAAGTCTC	629
QY	630	GCTGCAATCTGCGACGTGTGCACAAAGAAAGCGCTGCGAACAAGAAAGCTCTCGGGC	689
Db	3388	GCTGCAATCTGCGACGTGTGCACAAAGAAAGCGCTGCGAACAAGAAAGCTCTCGGGC	689
QY	690	CGGAGAGGGGGGGGGGGCGGCGGCAACCCGGCGACGCTTTGGCATCGTGGACCTTCGC	749
Db	3448	CGGAGAGGGGGGGGGGGCGGCGGCAACCCGGCGACGCTTTGGCATCGTGGACCTTCGC	749
QY	750	GCGCGGCGCCAGCGTACACAGCGACCTCATGTACATCGCGCAAGAAAGCGACGCGCGAG	809

Db 3508 GCGCGGCCCCAGCGTACACAGCGACCTCATTGTACATCCCGGAGAAAGCCAGCGCCGCGAG 3567

Q7 810 CCAGGCCAAGAGCAAGAGGCGCTGCGTCATCAGCTAG 846

Db 3568 CCAGGCCAAGAGCAAGAGGCGCTGCGTCATCAGCTAG 3604

RESULT 13	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
AC073621	58882 bp	DNA	linear	HTG_19-AUG-2002			
				clone CTD-3073J20 map 17, *** SEQUENCING			
				IN PROGRESS *** , 2 ordered pieces.			
AC073621.11	GI:22297517						
HTG: HTGS_PHASE2, HTGS_FULLTOP, HTGS_ACTIVERIN.							
human.							
Human							
HTG: HTGS_PHASE2, HTGS_FULLTOP, HTGS_ACTIVERIN.							
human.							
Human							
HTG: HTGS_PHASE2, HTGS_FULLTOP, HTGS_ACTIVERIN.							
human.							
Human							
HTG: HTGS_PHASE2, HTGS_FULLTOP, HTGS_ACTIVERIN.							
human.							
Human							
HTG: HTGS_PHASE2, HTGS_FULLTOP, HTGS_ACTIVERIN.							
human.							
Human							
HTG: HTGS_PHASE2, HTGS_FULLTOP, HTGS_ACTIVERIN.							
human.							
Human							
HTG: HTGS_PHASE2, HTGS_FULLTOP, HTGS_ACTIVERIN.							
human.							
Human							
HTG: HTGS_PHASE2, HTGS_FULLTOP, HTGS_ACTIVERIN.							
human.							
Human							
HTG: HTGS_PHASE2, HTGS_FULLTOP, HTGS_ACTIVERIN.							
human.							
Human							
HTG: HTGS_PHASE2, HTGS_FULLTOP, HTGS_ACTIVERIN.							
human.							
Human							
HTG: HTGS_PHASE2, HTGS_FULLTOP, HTGS_ACTIVERIN.							
human.							
Human							
HTG: HTGS_PHASE2, HTGS_FULLTOP, HTGS_ACTIVERIN.							
human.							
Human							
HTG: HTGS_PHASE2, HTGS_FULLTOP, HTGS_ACTIVERIN.							
human.							
Human							
HTG: HTGS_PHASE2, HTGS_FULLTOP, HTGS_ACTIVERIN.							
human.							
Human							
HTG: HTGS_PHASE2, HTGS_FULLTOP, HTGS_ACTIVERIN.							
human.							
Human							
HTG: HTGS_PHASE2, HTGS_FULLTOP, HTGS_ACTIVERIN.							
human.							
Human							
HTG: HTGS_PHASE2, HTGS_FULLTOP, HTGS_ACTIVERIN.							
human.							
Human							
HTG: HTGS_PHASE2, HTGS_FULLTOP, HTGS_ACTIVERIN.							
human.							
Human							
HTG: HTGS_PHASE2, HTGS_FULLTOP, HTGS_ACTIVERIN.							
human.							
Human							
HTG: HTGS_PHASE2, HTGS_FULLTOP, HTGS_ACTIVERIN.							
human.							
Human							
HTG: HTGS_PHASE2, HTGS_FULLTOP, HTGS_ACTIVERIN.							
human.							
Human							
HTG: HTGS_PHASE2, HTGS_FULLTOP, HTGS_ACTIVERIN.							
human.							
Human							
HTG: HTGS_PHASE2, HTGS_FULLTOP, HTGS_ACTIVERIN.							
human.							
Human							

TITLE	Direct Submission
JOURNAL	Submitted (19-AUG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT	On Aug 19, 2002 this sequence was

Smilt, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>
 Genome Center

TITLES	TRAEVERS, M., TRAVIS, N., TRIGGLIO, J., VASSILIEV, H., VIEL, R., WILSON, B., WU, X., WYMAN, D., YE, W. J., YOUNG, C., ZAHNOUN, J., ZEMBEK, L., ZIMMER, A. and ZODY, M.
TITLE	Direct Submission
JOURNAL	Submitted (03-MAR-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE	3 (bases 1 to 183598)
AUTHORS	Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N.

•
•
•
•

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 27, 2002, 22:19:47 ; Search time 3136 Seconds
(without alignments)
2607.745 Million cell updates/sec

Title: US-09-709-103-2
Perfect score: 1449
Sequence: 1 MRLAMIKKMCPSDSELSIP.....IREKASAGQAKDKRCVYS 281

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xlp
-O=/cgn2.1/USPTO.spool/US09709103/runat_20122002.162405.805/app.query.fasta.1.455
-DB=GenEmbl -OFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=bloms62 -TRANS=human40.cdt -LIST=45
-DOALIGN=200 -THR_SCORE=pcr -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFM=pic -NOR=ext -HEA=SIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USBR=US09709103.@cgn2.1.1.3637.@runat_20122002.162405.805 -NCP=6 -ICPU=3
-NO_XLPHY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMECUT=120
-WARN_TIMECUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database:

GenEmbl: *
1: gb_ba: *
2: gb_hgt: *
3: gb_in: *
4: gb_cm: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vl: *
15: em_ba: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_mu: *
20: em_mu: *
21: em_or: *
22: em_ov: *
23: em_pat: *
24: em_ph: *
25: em_pl: *
26: em_ro: *
27: em_sts: *
28: em_un: *

29: em_vl: *
30: em_htg_hum: *
31: em_htg_inv: *
32: em_htg_other: *
33: em_htg_mus: *
34: em_htg_pln: *
35: em_htg_rnd: *
36: em_htg_mam: *
37: em_htg_vrt: *
38: em_sy: *
39: em_htgo_hum: *
40: em_htgo_mus: *
41: em_htgo_other: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query Length	DB ID	Description
1	1449	100.0	846	9 AF498923	AF498923 Homo sapi
2	1449	100.0	979	9 AF172846	AF172846 Homo sapi
3	1449	100.0	1187	9 AF153192	AF153192 Homo sapi
4	1449	100.0	1740	9 AF069506	AF069506 Homo sapi
5	1449	100.0	1758	9 BC018041	BC018041 Homo sapi
6	1436	99.1	1746	9 AF177335	AF177335 Homo sapi
7	1405.5	97.0	1612	10 BC034166	BC034166 Mus muscu
8	1405.5	97.0	1623	10 AF009246	AF009246 Mus muscu
9	1399.5	96.6	1616	10 AF238157	AF238157 Rattus no
10	1387	95.7	4990	9 AF222979	AF222979 Homo sapi
11	1387	95.7	18334	9 AC020558	AC020558 Homo sapi
12	1384	95.5	5141	9 AF262018	AF262018 Homo sapi
13	1370	94.5	179124	10 AL603710	AL603710 Mouse DNA
14	1370	94.5	24899	2 AC025509	AC025509 Mus muscu
15	1356	93.6	183598	2 AC090608	AC090608 Homo sapi
16	1355	93.5	58882	2 AC073621	AC073621 Homo sapi
17	985	60.0	162504	2 AC122995	AC122995 Rattus no
18	893	61.6	2832	6 AX393244	AX393244 Sequence
19	893	61.6	2832	6 AX79143	AX79143 Homo sapi
20	893	61.6	2973	6 AX393267	AX393267 Sequence
21	893	61.6	3058	9 BC013419	BC013419 Homo sapi
22	867	59.8	3020	6 AX393362	AX393362 Sequence
23	867	59.8	3469	10 AF134409	AF134409 Rattus no
24	790	54.5	2699	9 HSM803172	ALB31936 Homo sapi
25	526	36.3	114771	9 H5669D19	AL022334 Human DNA
26	512	35.3	211071	10 AC076974	AC076974 Mus muscu
27	359.5	24.8	181528	3 AC010562	AC010562 Drosophila
28	359.5	24.8	207684	2 AC018039	AC018039 Drosophila
29	359.5	24.8	285495	3 AE003560	AE003560 Drosophila
30	359	24.8	2294	10 BC026377	BC026377 Mus muscu
31	342	23.6	4412	9 AB062937	AB062937 Maraca fa
32	334	23.1	1875	9 BC008065	BC008065 Homo sapi
33	334	23.1	2505	9 AB076889	AB076889 Homo sapi
34	334	23.1	4167	6 AX477382	AX477382 Sequence
35	334	23.1	143299	9 AL353619	AL353619 Human DNA
36	334	23.1	143857	9 AC021583	AC021583 Homo sapi
37	334	23.1	199612	9 AC021582	AC021582 Homo sapi
38	326	22.5	597	9 AY056037	AY056037 Homo sapi
39	326	22.5	597	9 AY059641	AY059641 Homo sapi
40	326	22.5	774	9 AB076888	AB076888 Homo sapi
41	326	22.5	2827	9 AK096600	AK096600 Homo sapi
42	326	22.5	3391	9 BC030660	BC030660 Homo sapi
43	326	22.5	177540	9 AC006538	AC006538 Homo sapi
44	324	22.4	126801	2 AC091518	AC091518 Mus muscu
45	324	22.4	201320	2 AC073816	AC073816 Mus muscu

ALIGNMENTS

RESULT 1

AF498923
LOCUS AF498923 846 bp mRNA linear PRI 01-MAY-2002
DEFINITION Homo sapiens activator of G protein signaling (RASD1) mRNA,
complete cds.
ACCESSION AF498923
VERSION AF498923.1 GI:20379021
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
REFERENCE Pohl, H.L., Ikeda, S.R. and Aronstam, R.S.
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
TITLE Direct Submission
JOURNAL Submitted (05-APR-2002) cDNA Resource Center, Guthrie Research
Institute, One Guthrie Square, Sayre, PA 18840, USA
FEATURES
source
1..846
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="Brain"
1..846
/gene="RASD1"
/note="AGS1"
1..846
/gene="RASD1"
/codon_start=1
/product="activator of G protein signaling"
/protein_id="AA021071.1"
/db_xref="GI:20379022"
/translation="MKLAAMIKKMKCPSELSIPAKNCYRMVILIGSKVKTAIVSR
LTCRFEDATPTTIDEDFRKRYSTINGEYQIDLDTSNHPFPARRKSLITGDFILY
EQLNNDPSFEVQLRQQLDITKSLNKNKKEVDPLVTCGNKGDHDFREVDREI
EQLVGDPPORCAFEISAKKNSSLDQFRALFMAKLPSEMSPLDHRKVSQYCDVLI
KKALRNKLLRAGSGGGGDPDAFGIVAFARPVSVDLMTIRKASAGSQAQKRE
RCVTS"
BASE COUNT 176 a 288 c 252 g 130 t
ORIGIN
Alignment Scores:
Pred. No.: 1,42e-127 Length: 846
Score: 1449.00 Matches: 281
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0
US-09-709-103-2 (1-281) x AF498923 (1-846)
QY 1 MetlyslauaiaalametllelyslmetysproseraspsergluleuSerillePro 20
DB 1 ATGAAGCTGGCGCGCATATCAGAGAGATGTCGCCGACGACGACGACGATATCCCG 60
QY 21 AlalysancysyyrarmetvalilleuuglyserSerlyValglyThrAlaile 40
DB 61 GCCAAGACTGCTATCGCATGTCATCTCCGCTGCTCAAGGTGGGAAACGCCATC 120
QY 41 ValSerarphleuThrGlyArgphelugluaspAlaIyrrhrProthrilegluaSphe 60
DB 121 GTGTGCGCCTTCCTCCACCGCGCGCTTCGAGAGACGCTACAGCCCTACCATGAGACTTC 180
QY 61 HisArglyspheylrSerilleargglyluValIyrrglnleuaspIleleuaspThrSer 80
DB 181 CACGCGCAAGTCTACTCCATCCGCGGAGGTCTACACGCTGACATCTCGACACGCTCC 240
QY 81 GlyasnHisprophetioAlaMetarqArgleuSerilleuThrglyAspValPheile 100
DB 241 GCGAAGCCCGCGTCCCGCGCATGCGGCGCTCTCCACCTCCACAGAGAGAGCTTTTCATC 300
QY 101 LeuValPheSerleuaspAsnArgaspSerphelugluValgluArgleuArglnGln 120
DB 301 CTGGTGTTCAGTCTGAGAACCGCGACCTCTTCGAGAGAGGTGACGCGCTCAGCAGCAG 360
QY 121 IleleuaspThrlySerCysleuLysAsnLysThrlySGluasnValaspValProleu 140
DB 361 ATCTCGACACCAAGTCTGCTTCACAGAACCAAGAGAGAGAGAGAGAGAGAGAGAGAG 420
QY 141 ValIleCysglyAsnLysglyAspPheArgAspPheArgIyrrglnValaspGlnArgIle 160
DB 421 GTCATCTGCGGCAACAGAGGTGACCGGACTTCTACCGCGAGGTGACCGCGGAGATC 480
QY 161 GluIleuValglyaspAspProgluArgCysAlaIyrrphelugluSerAlaLyslys 180
DB 481 GAGCAGCTGGTGGGCGAGACCGCCACGCGCTGGCGCTGCTCAGATCTGCGCAAGAG 540
QY 181 AsnSerSerleuaspGlnMetPheArgAlaLeuPheAlaMetAlaLysleuProSerGlu 200
DB 541 AACGACAGCTGGACCGACGATGTTCCGCGCTCTTCCATGAGCCAAAGTGCACGAG 600
QY 201 MetSerProaspIleuHisArglyValSerValgluThrCysaspValIleuHisLyslys 220
DB 601 ATGAGCCCAAGCTGACCGCGAAGGTCTCGTCCAGTACTGCGACGCTGCGCAAGAG 660
QY 221 AlaLeuArgAsnLysLysleuLeuArgAlaIyrrgIyrylyglyIyrylyaspProgly 240
DB 661 GCGCTGCGGAACAAAGAGCTGCTGCGGCGCGGACGCGCGCGCGCGCGCGCGCGCG 720
QY 241 AspAlaPhecllyIleValAlaIyrrphelAlaArgArgProSerValHisSerAspIleuMet 260
DB 721 GACGCTTGGCATCGTGGCACCTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTCATG 780
QY 261 TyrIleArgIyrylyAspAlaSerAlaIyrrgIyrylyAspIyrylyAspIyrylyValIle 280
DB 781 TACATCCGCGAAGAGCGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTCATG 840
QY 281 Ser 281
DB 841 AGC 843
RESULT 2
AF172846
LOCUS AF172846 979 bp mRNA linear PRI 05-JAN-2000
DEFINITION Homo sapiens ras-related protein (DEXRAS1) mRNA, complete cds.
ACCESSION AF172846
VERSION AF172846.1 GI:6014488
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
REFERENCE Pohl, H.L., Ikeda, S.R. and Aronstam, R.S.
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
TITLE Direct Submission
JOURNAL Submitted (27-JUN-1999) Department of Cell Biology, University of
Alabama at Birmingham, 1670 University Blvd., Birmingham, AL 35294, USA
FEATURES
source
1..979
/organism="Homo sapiens"
/db_xref="taxon:9606"
1..979
/gene="DEXRAS1"
/note="similar to Mus musculus ras-related protein encoded
by GenBank Accession Number AF009246; regulated by
glucocorticoid hormone"
/codon_start=1
/product="ras-related protein"

/protein_id="AA01364.1"
 /db_xref="GI:6014489"
 /translation="MKLAAMIKKMPDSELSIPAKNCYRMVILGSSKVGKTAIVSRF
 LTGRFEDATPTIEDFHRKFSIRGEVYOLDIDTSGNHPFAMRRLSITGDVFLV
 FSLNDRSFEVORLROOILDTKSLKNKTKENVDPVLVIGKMGDDFREVDOREI
 EOLVGDPPORCAVEFISAKKNSLDOMFPAKIPSEMSPDILHKKVVOYCDVIL
 KKALRNKKLLRAGSGGGGDPDAFGIVAFPARPVSVDLMTIREKASAGQAKDKE
 RCVIS"

BASE COUNT 195 a 347 c 284 g 153 t
 ORIGIN

Alignment Scores:

Pred. No.: 1.69e-127 Length: 979
 Score: 1449.00 Matches: 261
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0

US-09-709-103-2 (1-281) x AF172846 (1-979)

QY 1 MetLysLeuAlaMetIleLysLysMetCysProSerAspSerGluLeuSerIlePro 20
 DB 64 ATGAAGCTGGCGCGATGATCAAGAAGATGTCGCCGAGACTCGGAGCTGATATCCCG 123
 QY 21 AlalysAsnCyStyArgMetValIleLeuGlySerSerLysValGlyLysThrAlaIle 40
 DB 124 GCCAAGAACTGCTATGCGCAAGTGCATCTCGCTCTCCCAAGGTGGCGCAAGCGCCATC 183
 QY 41 ValSerArgPheLeuThrGlyArgPheGluAspAlaIleThrProThrIleGluAspPhe 60
 DB 184 GTGTGGCGCTTCTCTACCGCGCCCTTCGAGAGCGCTTACACGCTTACCTCGAGACTTC 243
 QY 61 HisArgLysPheThrSerIleArgGlyValIleValIleGluAspIleLeuAspThrSer 80
 DB 244 CACCGAAGTTCATCTACATCCGCGGAGAGTCTACAGCTGACATCTCGAGACTTC 303
 QY 81 GlyAsnHisProPheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIle 100
 DB 304 GGCACACACCGCTTCCCGCGCATGCGCGCTCTCCATCTCCACAGAGACGTTTTCATC 363
 QY 101 LeuValPheSerLeuAspAsnArgAspSerPheGluGluValGluAspIleLeuArgGln 120
 DB 364 CTGGTGTCTGATGTGGACAAACCGCGACTCTTCGAGAGGTGAGGCGCTCAGCGCAG 423
 QY 121 IleLeuAspThrLysSerCysLeuLysAsnLysThrLysGluAsnValAspValProLeu 140
 DB 424 ATCTCGACACCAAGTCTTGGCTCCAGAACAAACCAAGAGAGCTGACGCGCCCTG 483
 QY 141 ValIleCysGlyAsnLysGlyAspArgAspPheThrArgGluValAspGluIle 160
 DB 484 GTATCTGCGCGCAAGAGGTGACCGGACTTCTACCGGAGGTGACCGCGCGCATC 543
 QY 161 GluGluLeuValGlyAspAspProGlnArgCysAlaIleThrPheGluIleSerIleLys 180
 DB 544 GAGCAGCTGGTGGGACGACACCGCGCGCTGCTACTTCAAGATCTCGGCGCAAGAG 603
 QY 181 AsnSerSerLeuAspGlnMetPheArgAlaLeuPheAlaMetAlaLysLeuProSerGlu 200
 DB 604 AACAGAGAGCTGGACGAGATGTTCCGCGGCTCTTCCCATGCGCAAGCTGCCCGACGAG 663
 QY 201 MetSerProAspLeuHisArgLysValSerValGlnIleThrCysAspValLeuHisLys 220
 DB 664 ATGAGCCCAAGCTGCAACGCAAGGTCTGCTGCTACTTGGACGCTGTCGACAGAGAG 723
 QY 221 AlalysLeuArgAsnLysLysLeuLeuArgAlaGlySerGlyGlyGlyLysLysAspProGly 240
 DB 724 GCGCTCGGGAACAAGAAGCTGCTGGGCGGAGAGGCGGCGGCGCGCGCGCGCGCGG 783
 QY 241 AspAlaPheGlyIleValAlaProPheAlaArgArgProSerValHisSerAspLeuMet 260
 DB 784 GAGCGCTTGGCATCTGTCGACCGCTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTCAG 843

QY 261 TyrIleArgGluLysAlaSerAlaGlySerGlnAlaLysAspLysGluArgCysValIle 280
 DB 844 TACATCCGCGAGAGAGCGCCGCGGACCGCGAGCGCAAGACAGCGCGCTGCTATC 903
 QY 281 Ser 281
 DB 904 AGC 906

RESULT 3 AF153192 1187 bp mRNA linear PRI 02-JUN-1999
 LOCUS AF153192
 DEFINITION Homo sapiens ras-related protein mRNA, complete cds.
 ACCESSION AF153192
 VERSION AF153192.1 GI:4960166
 KEYWORDS
 SOURCE
 ORGANISM Homo sapiens.
 Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 1187)
 AUTHORS Kempainen,R.J.
 TITLE Identification of human pituitary Dextral
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1187)
 AUTHORS Kempainen,R.J.
 TITLE Direct Submission
 JOURNAL Submitted (20-MAY-1999) Anatomy & Physiology, Auburn University,
 College of Veterinary Medicine, Auburn, AL 36849, USA

FEATURES
 source

1. 1187
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="2"
 /map="2q32"
 /tissue-type="pituitary gland"
 1. 846
 /note="Dextral; similar to Mus musculus ras-related
 protein encoded by GenBank Accession Number AF009246"
 /codon_start=1
 /product="ras-related protein"
 /protein_id="AAD34621.1"

US-09-709-103-2 (1-281) x AF153192 (1-1187)
 BASE COUNT 249 a 396 c 362 g 180 t
 ORIGIN

Alignment Scores:

Pred. No.: 2.13e-127 Length: 1187
 Score: 1449.00 Matches: 281
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0

QY 1 MetLysLeuAlaMetIleLysLysMetCysProSerAspSerGluLeuSerIlePro 20
 DB 1 ATGAAGCTGGCGCGATGATCAAGAAGATGTCGCCGAGACTCGGAGCTGATATCCCG 60
 QY 21 AlalysAsnCyStyArgMetValIleLeuGlySerSerLysValGlyLysThrAlaIle 40
 DB 61 GCCAAGAACTGCTATGCGCAAGTGCATCTCGCTCTCCCAAGGTGGCGCAAGCGCCATC 120
 QY 41 ValSerArgPheLeuThrGlyArgPheGluAspAlaIleThrProThrIleGluAspPhe 60
 DB 121 GTGTGGCGCTTCTCTACCGCGCCCTTCGAGAGCGCTTACACGCTTACGAGACTTC 180

QY 61 HisArgLysPheTyrSerIleArgIleValTyrGlnLeuAspIleLeuAspThrSer 80
 DB 181 CACGCGAAGTTCTACTCCATCCGCGGAGGTCTACACAGCTCGACATCTCGACAGCTCC 240
 QY 81 GlyAsnHisProPheProAlaMetArgIleuSerIleLeuThrGlyAspValPheIle 100
 DB 241 GGCACACACCGCTTCCCGCATCGGCGCTCTCCATCTCCACAGAGACGTTTTCATC 300
 QY 101 LeuValPheSerIleuAspAsnArgAspSerPheGlnIleValGlnArgLeuArgGln 120
 DB 301 CTGGTGTTCACCTTGACCAACCGGACTCTCTCAGAGAGGTGCGCGCTCAGGACAG 360
 QY 121 IleLeuAspThrLysSerCysLeuLysAsnLysThrLysGlnAsnValAspValProLeu 140
 DB 361 ATCTCGACACCAAGTCTTGCTCAAGAACCAAGGAGAACGTTGAGCTGCCCTG 420
 QY 141 ValIleCysGlnAsnLysGlyAspArgAspPheTyrArgGlnValAspGlnArgGlnIle 160
 DB 421 GTCACTGCGGCAACAAAGGGGACCGGACTTTCACGCGAGGTGACACAGCGAGATC 480
 QY 161 GluGlnLeuValGlyAspAspProGlnArgCysAlaTyrPheGlnIleSerAlaLysLys 180
 DB 481 GAGCAGCTGGTGGGCGAGACCGCCAGCGCTGCTGAGATCTCGGCGCAAGAG 540
 QY 181 AsnSerSerLeuAspGlnMetPheArgAlaLeuPheAlaLysLysLeuProSerGln 200
 DB 541 AACAGCAGCCCTGACACAGATGTCTCCGCGCTTCCGCAATGCCAAGCTCCACAGCAG 600
 QY 201 MetSerProAspLeuHisArgLysValSerValGlnTyrCysAspValLeuHisLysLys 220
 DB 601 ATGAGCCCGAGCTGACCGCGAAGGTCTCGGTGCGAGTCTGCGAGCTGCGCAAGAG 660
 QY 221 AlaLeuArgAsnLysLysLeuLeuArgAlaGlySerGlnGlyGlyGlyAspProGly 240
 DB 661 GCGCTGGCGAACAAGAGCTGCTGCGGCGCGAGCGGCGGCGGCGGAGACCGCGGC 720
 QY 241 AspAlaPheGlyIleValAlaProPheAlaArgArgProSerValHisSerAspLeuMet 260
 DB 721 GACCGCTTGGCATCTGCGCACCTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 780
 QY 261 TyrIleArgIleLysAlaSerAlaGlySerGlnAlaLysAspLysGlyLysValIle 280
 DB 781 TACATCCGCGAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGC 840
 QY 281 Ser 281
 DB 841 AGC 843

RESULT 4
 AF069506 1740 bp mRNA linear PRI 10-JAN-2000
 LOCUS Homo sapiens activator of G protein signaling (AGS1) mRNA, complete cds.
 DEFINITION
 ACCESSION AF069506
 VERSION AF069506.1 GI:4959037
 KEYWORDS
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 1740)
 Cismowski,M.J., Takesono,A., Ma,C., Lizzano,J.S., Xie,X., Fuernkrantz,H., Lanier,S.M. and Duzic,E.
 Genetic screens in yeast to identify mammalian nonreceptor modulators of G-protein signaling
 Nat. Biotechnol. 17 (9), 878-883 (1999)
 JOURNAL MEDLINE
 PUBMED 9940338
 10471929
 2 (bases 1 to 1740)
 Cismowski,M.J., Fuernkrantz,H., Ma,C., Spruyt,M., Xie,X., Lanier,S.M. and Duzic,E.
 Direct Submision
 TITLE Submitted (01-JUN-1998) Receptor Pharmacology/Biochemistry, Cadus

Pharmaceutical Corporation, 777 Old Saw Mill River Rd., Tarrytown, NY 10591, USA
 FEATURES
 source 1. 1740
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /tissue_type="liver"
 /dev_stage="adult"
 1. 1740
 /gene="AGS1"
 146. 991
 /note="AGS1"
 /note="ras-related G protein; similar to Mus musculus product encoded by Genbank Accession Number AF009246"
 /codon_start=1
 /product="activator of G protein signaling"
 /protein_id="ABD34206.1"
 /db_xref="GI:4959038"
 /translation="MKLAMIRKMCPSDSELSIPAKNCYRMVILGSSKVGKTAIVSRF LTGREDAVTPRIEDFHRKFFYSIRGEVYQDLDTSGNHPPARRRSLITGDFILY FSLDRDSEEEVORLQOILPTKSLKRTKENVDPLVIGNGRDRFYREVQREI EQLVGDDPORCAVEFEISAKNSISLDMERALFPAKXLPSESPDLHRRVSVQYCDVH KKALNKMLNRAGSGGGGDPGDAGTAPPARRRPSVSDIMTIRKASGSAKKKE KCVIS"
 BASE COUNT 422 a 546 c 483 g 289 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 3.37e-127 Length: 1740
 Score: 1449.00 Matches: 281
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0
 US-09-709-103-2 (1-281) x AF069506 (1-1740)
 QY 1 MetLysLeuAlaAlaMetIleLysLysMetCysProSerAspSerGlnLeuSerIlePro 20
 DB 146 ATGAAACTGGCCCGATGATCAAGAGATGTCGCCAGCGACTCGAGCTGAGTATCCG 205
 QY 21 AlaLysAsnCysTyrArgMetValIleLeuGlySerSerValGlyLysThrAlaIle 40
 DB 206 GCCAAGAACTGCTATGCAATGTCATCTGCTGCTCAAGGTGGGCAAGAGCGCATC 265
 QY 41 ValSerArgPheLeuThrGlyArgPheGlnAspAlaTyrThrProThrIleGlnAspPhe 60
 DB 266 GTGTGGGCTTCTTCAACGCGCGCTTCAAGAGCGCTTACAGCCTACAGCTGAGACTTC 325
 QY 61 HisArgLysPheTyrSerIleArgIleValTyrGlnLeuAspIleLeuAspThrSer 80
 DB 326 CACCGCAAGTCTACATCCCGCGCGAGGCTTACACCTCGACATCTCGACAGCTCC 385
 QY 81 GlyAsnHisProPheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIle 100
 DB 386 GGCACACCGCTTCCCGCATCGGCGCTTCTCATCTCAGACAGAGCGTTTTCATC 445
 QY 101 LeuValPheSerLeuAspAsnArgAspSerPheGlnIleValGlnArgLeuArgGln 120
 DB 446 CTGGTGTTCAGTGTGCAACACCGCGACTCTTCAAGAGGTGAGGCGTCAAGCAGCG 505
 QY 121 IleLeuAspThrLysSerCysLeuLysAsnLysThrLysGlnAsnValAspValProLeu 140
 DB 506 ATCTCGACACCAAGTCTTGCTCAAGAACCAAGGAGAACGTGAGCTGCCCTG 565
 QY 141 ValIleCysGlnAsnLysGlyAspArgAspPheTyrArgGlnValAspGlnArgGlnIle 160
 DB 566 GTCACTGCGGCAACAAAGGTGACCGGACTTTCACGCGAGGTGACGCGAGATC 625
 QY 161 GluGlnLeuValGlyAspAspProGlnArgCysAlaTyrPheGlnIleSerAlaLysLys 180
 DB 626 GAGCAGCTGGTGGGCGAGACCGCCAGCGCTGCTTCTGAGATCTCGGCGCAAGAG 685

QY 181 AsnSerSerLeuAspGlnMetPheArgAlaLeuPheAlaMetAlaLysLeuProSerGlu 200
 DB 686 AACAGACGCTGGACGAGATGTTCCGGCGCTCTTCCGATGGCCAAAGTCCCGACGAG 745
 QY 201 MetSerProAspLeuHisArgLysValSerValGlnTyrCysAspAlaLeuHisLysLys 220
 DB 746 ATAGAGCCAGACTGCACCGCAAGGTCTGGTCACTACTGCGACCTGCTGCACAAAGAG 805
 QY 221 AlaLeuArgAsnLysLysLeuLeuArgAlaGlySerGlyGlyGlyGlyGlyGlyGly 240
 DB 806 GCGCTGGCAAGAAAGAGTGTGCGGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 865
 QY 241 AspAlaPheGlyLysValAlaPheProPheAlaArgProSerValHisSerAspLeuMet 260
 DB 866 GAGCGCTTGGCATCGTCGCGACCTTCCGCCCGCGCGCGCGCGCGCGCGCGCGCGCG 925
 QY 261 TyrIleArgGlyLysValSerAlaGlySerGlnAlaLysAspLysGluArgCysValIle 280
 DB 926 TACATCCGAGAGAGCG 985
 QY 281 Ser 281
 DB 986 AGC 988

RESULT 5
 BC018041 1758 bp mRNA linear PRI 06-DEC-2001
 LOCUS
 DEFINITION Homo sapiens, RAS, dexamethasone-induced 1, clone MGC:26290
 IMAGE:497482, mRNA, complete cds.
 ACCESSION BC018041
 VERSION BC018041.1 GI:17390075
 KEYWORDS MGC.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens.
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1758)
 Strausberg, R.
 Direct Submission
 Submitted (03-DEC-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
 Contact: MGC help desk
 Email: cgabs@email.nih.gov
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiroki
 Toshiyuki and Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome
 Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www.shgc.stanford.edu>
 Contact: (Dickson, Mark) modpaxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
 R. M.

REMARK
 COMMENT
 Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
 Series: IRAC Plate: 32 Row: f Column: 11.

FEATURES

source

Location/Qualifiers

1..1758

/organism="Homo sapiens"

/db_xref="taxon:9606"

/db_xref="taxon:9606"

/clone="MGC:26290 IMAGE:497482"

/tissue_type="Brain, hypothalamus"

/clone_id="NH_MGC_96"

/lab_host="DH10B"

/note="Vector: pBluescript"

213..1058

/codon_start=1

/product="RAS, dexamethasone-induced 1"

CDS

/product="RAS, dexamethasone-induced 1"

/protein_id="AAH18041.1"
 /db_xref="GI:17390075"
 /translation="MKLAAMTKKCPDSELSIPAKCYRMVILGSKVKTAIVSR
 LTRGRFADYPTIEDFHRKREYSIGEVYQIDLDTSNHPFPARRRLSLTGVFLY
 FSLNRSFEEVQRLRODLDYTKSCLNKRKENVDYLVKNGSDGFREVDQREI
 EQLYGDPORCAYFEISAKKNSLIDQKFAAKLPSESPDLHKVSVQCDVH
 KKALRNKLLRAGSGGGGDPGDAFGIVAPFARRPSVHSDLMTREKASGSAOKDE
 RCVIS"

BASE COUNT 387 a 576 c 501 g 294 t
 ORIGIN

Alignment Scores:
 Pred. No.: 3,42e-127 Length: 1758
 Score: 1449.00 Matches: 281
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: Gaps: 0

US-09-709-103-2 (1-281) x BC018041 (1-1758)

QY 1 MetLysLeuAlaAlaMetIleLysLysMetCysProSerAspSerGluLeuSerIlePro 20
 DB 213 ATGAAGACCTGGCGCGCATGATCAAGAAAGATGTCGCGAGGAGTGGAGCTGATCCG 272
 QY 21 AlaLysAsnCysTyrArgMetValIleLeuGlySerSerLysValGlyLysThralle 40
 DB 273 GCCAAGACGTCATGCGATGTCATCCGCGCTGTCACAGGTGGCGGCGGCGCGCATC 332
 QY 41 ValSerArgPheLeuThrCylarPheGluAspAlaTyrThrProThrIleGluAspHe 60
 DB 333 GTTCGCGCTTCTCTACCGCGCGCTTCCAGAGCGCTACAGCGCTTACCACTGAGGACTTC 392
 QY 61 HisArgLysPheTyrSerIleArgGlyGluValTyrGlnLeuAspIleLeuAspPheSer 80
 DB 393 CACCGCAATTCATCCATCCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 452
 QY 81 GlyAsnHisProPheProAlaMetArgLysSerIleLeuThrGlyAspValPheIle 100
 DB 453 GGGACACACCGCTTCCCGCCCATGCGCGCGCTCTCATCTCCACAGAGACCTTTTCATC 512
 QY 101 LeuValPheSerLeuAspAsnArgAspSerPheGluGluValAlaArgLeuArgGln 120
 DB 513 CTGGTTCATGCTGCACACCGCGCATCTCTTCGAGAGGTGACGCGCTCAGGACGAC 572
 QY 121 IleLeuAspThrLysSerCysLeuLysAsnLysTyrLysGluAsnValAspValProLeu 140
 DB 573 ATCTTGACACCAAGATCTTGCCTCAAGAACAAACCAAGAGAGAGTGGAGCTGCCGCG 632
 QY 141 ValIleCysGlyAsnLysGlyAspArgAspPheTyrArgGluValAlaAspGlnArgLys 160
 DB 633 GTTCATCTGGCGCAACAGGAGTACCGGCTTCTACCGGAGGTGACGCGGCGAGATC 692
 QY 161 GluGluLeuValGlyAspAspProGlnArgCysAlaTyrPheGlnIleSerAlaLysLys 180
 DB 693 GAGCAGCTGGTGGCGACACACCGCGCGCTGCGCTTCTGAGATCTCGGCGCAAGAG 752
 QY 181 AsnSerSerLeuAspGlnMetPheArgAlaLeuPheAlaMetAlaLysLeuProSerGlu 200
 DB 753 AACAGACGCTGGACGAGATGTTCCGGCGCTCTTCCGATGGCCAAAGTCCCGACGAG 812
 QY 201 MetSerProAspLeuHisArgLysValSerValGlnTyrCysAspAlaLeuHisLysLys 220
 DB 813 ATAGAGCCAGACTGCACCGCAAGGTCTGGTCACTACTGCGACGCTGCGACAGAGAG 872
 QY 221 AlaLeuArgAsnLysLysLeuLeuArgAlaGlySerGlyGlyGlyGlyGlyGlyGly 240
 DB 873 GCGCTGGCAAGAAAGTGTGCGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 932
 QY 241 AspAlaPheGlyLysValAlaPheProPheAlaArgProSerValHisSerAspLeuMet 260
 DB 933 GAGCGCTTGGCATCGTCGCGACCTTCCGCCCGCGCGCGCGCGCGCGCGCGCGCGCG 992

```

OY 261 TTTTleatggtulysalaserAlaglySerGlnAlalysAspLysgUatgCysValIle 280
DB 993 TACATCCCGAGAGAGCCGCGCGGAGCCAGGCAAGAGAGAGAGCGCTGGCTATC 1052
OY 281 Ser 281
DB 1053 AGC 1055

RESULT 6
AF177335 1746 bp mRNA linear PRI 03-OCT-2000
LOCUS Homo sapiens clone SP1942 unknown mRNA.
DEFINITION AF177335
ACCESSION AF177335.1 GI:10503968
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 1746)
Gu,J.R., Wan,D.F., Zhao,X.T., Zhou,X.M., Jiang,H.Q., Zhang,P.P.,
Qin,W.X., Huang,Y., Qiu,X.K., Qian,L.F., He,L.P., Li,H.N., Yu,Y.,
Yu,J. and Han,L.H.
Novel human cDNA clone with function of inhibiting cancer cell
growth
TITLE
Unpublished
2 (bases 1 to 1746)
Gu,J.R., Wan,D.F., Zhao,X.T., Zhou,X.M., Jiang,H.Q., Zhang,P.P.,
Qin,W.X., Huang,Y., Qiu,X.K., Qian,L.F., He,L.P., Li,H.N., Yu,Y.,
Yu,J. and Han,L.H.
Direct Submission
JOURNAL
Submitted (06-AUG-1999) National Laboratory For Oncogenes & Related
Genes, Shanghai Cancer Institute, 25/Jn 2200 Xie Tu Road, Shanghai
200032, P.R. China
FEATURES
source
1..1746
location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="SP1942"
455..1039
/codon_start=1
/product="unknown"
/protein_id="AAG17979.1"
/db_xref="GI:10503968"
/translacion="MRRSLILGVDFILVFSLDNDSFEVRLRQILITKSLKMK
TKENVPLVIGNGKRDYREVDREIQVGDPOPCAFETISAKNSISIDQMR
ALFAMAKLPSEMSPLHRYVYQYCDVILKALKLRKLLRAGSGGGGDPDAFSTVA
PFARPSVSDIMYIRKASAGSOAKDKRCVTS"
BASE COUNT 388 a 569 c 496 g 293 t
ORIGIN
Alignment Scores:
Pred. No.: 5.7e-126 Length: 1746
Score: 1436.00 Matches: 281
Percent Similarity: 99.65% Conservative: 0
Best Local Similarity: 99.65% Mismatches: 0
Query Match: 99.10% Indels: 1
Gaps: 0

US-09-709-103-2 (1-281) x AF177335 (1-1746)
OY 1 MetLysleuAlaMetIleLysMetCysProSerAspSerGluLeuSerIlePro 20
DB 193 ATGAAACTGGCCCGCATGATCATCAGAGATGTCGCCAGGAGACTGGAGCTATCCCG 252
OY 21 AlalysAsnCySTyRarMetValIleLeuGlySerSerLysValGlyLysThrAlaIle 40
DB 253 GCCAAGAACTGATCGATCGATCGCTCGCTCCCAAGGTGGGCAAGAGCGCATC 312
OY 41 ValSerArgPheLeuThrGlyArgPheGluAspAlaTyThrProThrIleGluAspPhe 60
DB 313 GTGTGGCGTTCTCTACCGCGCTTCGAGAGCGCTACAGCGCTACCATCGAGGACTTC 372

OY 61 HleArgLysPheThrSerIleArgGlyGluValTyArgLleu-AspIleLeuAspThrSe 80
DB 373 CACCGCAAGTCTACTCCATCCCGCGGAGCTTACCACTTCGACATCTCGACACGTC 432
OY 80 rGlyAsnHisProPheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheI 100
DB 433 CGGCAACACACCGCTCCCGCCATGCGGCGCTCTCCATCTCACAGAGAGGTTTCAT 492
OY 100 eleuValPheSerLeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGlnI 120
DB 493 CTTGTGTTCAGTCTGACACACCGGACCTCTTCGAGAGGAGGAGCGGCTTCAGACACA 552
OY 120 nIleLeuAspThrLysSerCysLeuLysAsnLysThrLysGluAsnValAspValProIe 140
DB 553 GATCTCGACACCAAGTCTTGCTTCAGAAACAAACAGAGAAACGTGGAGCTCCCTT 612
OY 140 uValIleCysGlyAsnLysGlyAspArgAspPheTyArgGluValAspGlnArgGluI 160
DB 613 GGTCACTCGCGGACACAGAGGTGACCGGAGCTTACCGCGAGGTGACACGCGGAGAT 672
OY 160 eGluGlnLeuValGlyAspAspProGlnArgCysAlaTyArgPheGluIleSerAlaLysLy 180
DB 673 CGAGCAGCTGTGGGCGGAGACACCGGCGCTGCGCTTACCTTCGAGATCTCGCCAAAGA 732
OY 180 sAsnSerSerLeuAspGlnMetPheArgAlaLeuPheAlaMetAlaLysLeuProSerGI 200
DB 733 GAACAGCAGCCTTGAGACACAGATGTCGCGGCTCTTCGCGCAGGCGCAGCTCCAGCGA 792
OY 200 uMetSerProAspLeuHisArgLysValSerValGlnTyrcysAspValLeuHisLysLy 220
DB 793 GATGAGCCCGACGCTGACCGGACGAGTCTCGGTGACGACTGCGAGCTGCGACAAAGA 852
OY 220 sAlaLeuArgAsnLysLysLeuLeuArgAlaGlySerGlyGlyGlyGlyAspProGI 240
DB 853 GCGCGTGGGAGAACAGAAAGACTCTGCGGCGCGGACGCGGCGGCGGCGGCGGCGG 912
OY 240 yAspAlaPheGlyIleValAlaProPheAlaArgArgProSerValHisSerAspLeuMe 260
DB 913 CGACGCTTTGGCAGTCGTCGACCTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTCAT 972
OY 260 tTyrlleatggtulysalaserAlaglySerGlnAlalysAspLysgUatgCysValI 280
DB 973 GTACATCCCGAGAGAGCCAGCGCGGAGCCAGGCAAGAGAGAGAGCGCTGGCTAT 1032
OY 280 eSer 281
DB 1033 CAGC 1036

RESULT 7
BC034166 1612 bp mRNA linear ROD 07-AUG-2002
LOCUS BC034166
DEFINITION Mus musculus, RAS, dexamethasone-induced 1, clone MGC:56188
IMAGE:4989312, mRNA, complete cds.
ACCESSION BC034166
VERSION BC034166.1 GI:21706874
KEYWORDS MGC.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 1612)
Strausberg,R.
Direct Submission
Submitted (02-JUL-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2550,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
REMARK
COMMENT

```

DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www-shgc.stanford.edu>
 Contact: (Dickson, Mark) mcdickax1.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAC Plate: 58 Row: 9 Column: 5
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 6677672.

FEATURES

source

1. 1612
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /map="FVB/N"
 /clone="MGC:36188 IMAGE:4989112"
 /tissue_type="Colon, normal, 5 month old male mouse."
 /clone_lib="NCI_CGAP_Co24"
 /lab_host="DH10B"
 /note="Vector: PCMV-SPORT6"
 108. 950
 /codon_start=-1
 /product="RAS, dexamethasone-induced 1"
 /protein_id="AAH34166.1"
 /db_xref="GI:21706875"
 /db_xref="LocusID:19416"
 /translation="MTLAKMKRMPDSELSIPAKNCYRMVILGSSKVKATVSRF
 LTRPRDATPTIEDPRKRFYIRGVEYDILDLISGNHPRPMRLSLILDDVILY
 FSLNDRDSFEVQRKQOILIDTRSLKNTKENVDPVLTGKNGDRDYREVQREI
 EQLVGDPPQRCAYFEISAKNSLDQMFALPAMKLPSEMSPDILRKYSVQCVLH
 KRALRNKLLRAGSGGGGHDGAFILAFARPSVHSDLMYIREKTSVQAKDKER
 CVYS"

CDS

BASE COUNT 417 a 424 c 420 g 351 t
 ORIGIN

Alignment Scores:

Pred. No.: 3.88e-123 Length: 1612
 Score: 1405.50 Matches: 274
 Percent Similarity: 98.58% Conservative: 3
 Best Local Similarity: 97.51% Mismatches: 1
 Query Match: 97.00% Indels: 1
 DB: 10 Gaps: 1

US-09-709-103-2 (1-281) x BC034166 (1-1612)

QY 1 MetLysLeuAlaMetLysLysMetCysProSerAspSerGluLeuSerIlePro 20
 Db 108 ATGAACCTGCGCGGATGATCAAGAAATGTCGCCAAGCGACCTGGAACAGATATCCCG 167
 QY 21 AlalysAsnCysTyrArgMetValIleLeuGlySerSerLysValGlyLysThrAlaIle 40
 Db 168 GCCAAGACAGCTACAGAGATGCTATCTCGGCTCATCCAAAGGGGCAAGACGCCCAT 227
 QY 41 ValSerArgPheLeuThrGlyArgPheGluAspAlaTyrThrProThrIleGluAspPhe 60
 Db 228 GTGTCGGGCTCTCCACGGCGGCTTTCGAGGATGCTTACACCCCTACACATCGAGAGACTTC 287
 QY 61 HisArgLysPheTyrSerIleArgGlyGluValTyrGlnLeuAspIleLeuAspPheSer 80
 Db 288 CACGAAAGTTTACTGATCCGCGGGAAGTCAACAGTTGACATCTGACACATCC 347
 QY 81 GlyAsnHisProPheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIle 100
 Db 348 GGCATATCATCGTTTCCCGCATCGCGGCTCTCTATCTCTACACAGAGAGAGCTTTTATT 407
 QY 101 LeuValPheSerLeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGlnGln 120
 Db 408 CTGGTGTTCACCTTAGAACACCGGAGCTCATTCGAAAGAGGTCAAAAGCTCAACAGACAG 467
 QY 121 IleLeuAspThrLysSerCysLeuLysAsnLysThrLysGluAsnValAspValProLeu 140

Db 468 ATCTAGACACCACTCTCTCTCAAGAACAAACCAAGAGATGTCAGCTGCCCTG 527
 QY 141 ValIleCysGlyAsnLysGlyAspArgAspPheTyrArgGluValAspGlnArgGluIle 160
 Db 528 GTCATTGGGGTAACAAAGGGGACCGGGACTTCTACCGGGGAAGTAGACACGGGAGATT 587
 QY 161 GluGlnLeuValGlyAspAspProGlnArgCysAlaTyrPheGluIleSerAlaLysLys 180
 Db 588 GAGCAGCTGGTGGGAGACACCTTCAGCGTGTGCTACTTCTGAGATCTCAGCCAGAGAG 647
 QY 181 AsnSerLeuAspGlnPheArgAlaLeuPheAlaMetAlaLysLeuProSerGlu 200
 Db 648 AACAGAGCTTGGACCAAGATGTTCCGTGGCTCTTGGCATAGGCGCAAGCTGCTAGGCGAG 707
 QY 201 MetSerProAspLeuHisArgLysValSerValGlnTyrCysAspValLeuHisLysLys 220
 Db 708 ATGAGCCCGGATTCGACCGCAAGATCTGTGCGAGTACTGGAGCTACTGACACAGAGAG 767
 QY 221 AlalysAsnLysLysLysLeuArgAlaGlySerGlyGlyGlyGlyAspProGly 240
 Db 768 GCTGTAGAGAACAAAGAGCTTCTGCGTGGCGGACG--GGAGGGGGGGGAGCACCGGC 824
 QY 241 AspAlaPheGlyIleValAlaProPheAlaArgArgProSerValHisSerAspLeuMet 260
 Db 825 GATGCTTGGCATCTTGCGCGCTTGGCTGCGACACCGAGGTCACACGACCTCATG 884
 QY 261 TyrIleArgGluAlaAlaSerAlaGlySerGlnAlaLysAspLysGluArgCysValIle 280
 Db 885 TCATTCGCGGAAAAAACAGCTGCGGACCGACGAGGTAAAGAACAGACGGCTGTATC 944
 QY 281 Ser 281
 Db 945 AGT 947

RESULT 8
 AF009246
 LOCUS 1623 bp mRNA linear ROD 13-FEB-1998
 DEFINITION Mus musculus ras-related protein (DEXRAS1) mRNA, complete cds.
 ACCESSION AF009246
 VERSION AF009246.1 GI:2253712
 KEYWORDS
 SOURCE
 ORGANISM

Mus musculus.
 Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 Dexamethasone rapidly induces a novel ras superfamily
 member-related gene in Atr-20 cells
 J. Biol. Chem. 273 (6), 3129-3131 (1998)

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL

2 (bases 1 to 1623)
 Kempainen, R.J. and Behrend, E.N.
 Direct Submission
 Submitted (18-JUN-1997) Physiology & Pharmacology, Auburn
 University College of Veterinary Medicine, 213 Greene Hall, Auburn,
 AL 36849, USA

FEATURES

source

1. 1623
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /cell_line="corticotrope tumor cell line Atr-20"
 1. 1623
 /gene="DEXRAS1"
 142. 984
 /gene="DEXRAS1"
 /note="Induced by dexamethasone"
 /codon_start=-1
 /product="ras-related protein"
 /protein_id="AAC5338.1"
 /db_xref="GI:2253713"
 /translation="MTLAKMKRMPDSELSIPAKNCYRMVILGSSKVKATVSRF

gene

CDS

ITGREDATPTTIEDHFRKFSYIRGEVYOLDITDSCNHPBPAMRLSITGDVILY
 FSLDNRDSFEVQRLKQOIIDTKSLNKRKENDVPLVTCGNKGDYFREVQREI
 EQLVDDPQRCATFEISAKRNSLDQRFALPSEMSPDILHRKVSQYCDVILH
 KRALNRKRLLRAGSGGGHDGDAFGIILAPARRPSVSHDLMYIREKTSVSOAKDKER
 CVIS"

BASE COUNT 412 a 437 c 417 g 357 t
 ORIGIN

Alignment Scores:

Pred. No.: 3,91e-123 Length: 1623
 Score: 1405.50 Matches: 274
 Percent Similarity: 98.58% Conservative: 3
 Best Local Similarity: 97.51% Mismatches: 3
 Query Match: 97.00% Indels: 1
 DB: 10 Gaps: 1

US-09-709-103-2 (1-281) x AF009246 (1-1623)

QY 1 MetLysLeuAlaAlaMetIleLysLysMetCysProSerAspSerGluLeuSerIlePro 20
 DB 142 ATGAAGTGGCCGGCATGATCAAGAAAGTGGCCCAACCGCATCTGAACTGATATCCG 201
 QY 21 AlaLysAsnCysTyrArgMetValIleLeuGlySerSerLysValGlyLysThrAlaIle 40
 DB 202 GCCAAGAACTGCTACAGATGATGATCTCCGCTCATCCAAAGTGGCAGAGCCCATTT 261
 QY 41 ValSerArgPheLeuThrGlyArgPheGluAspAlaTyrThrProThrIleGluAspPhe 60
 DB 262 GTGTGCGCCCTTCCTCCACGGCGGCTTTCGAGGATGCTTACCCCTACCATGAGACTTC 321
 QY 61 HisArgLysPheTyrSerIleArgGlyGluValTyrGlnLeuAspIleLeuAspThrSer 80
 DB 322 CACGAAAGTTTACTGATCCGGCGGAAGTCAACAGTTGGACATAGTGCACATCC 381
 QY 81 GlyAsnHisProPheProAlaMetLargArgLeuSerIleLeuThrGlyAspValPheIle 100
 DB 382 GGCAATATCCGTTCCCGGCATGGCGCCCTCTATCCCTACAGAGAGAGCTTTTATT 441
 QY 101 LeuValPheSerLeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGlnGln 120
 DB 442 CTGGTGTTCACCTTACACMACCGGACTCATTCGAAAGGTGCAAGGCTCAACACAGCAG 501
 QY 121 IleLeuAspThrLysSerCysLeuLysAsnLysThrLysGluAsnValAspValProLeu 140
 DB 502 ATCTTAGACACCAAGTCTGCTCAAGAAACAAACCAAGAAATGGAGCGTCCCTG 561
 QY 141 ValIleCysGlyAsnLysGlyAspArgAspPheTyrArgGluValAspGlnArgGlnIle 160
 DB 562 GTCAATTCGCGTAAACAAGGGGACCGGAGACTTTCACCGGGAAGTACAGCAGCGGAGATT 621
 QY 161 GluGlnLeuValGlyAspAspProGlnArgCysAlaTyrPheGluIleSerAlaLysLys 180
 DB 622 GAGAGCTGTGGTGGACGACCCCTACGCTTGTCTCCACTTCGAGATCTCAGCCAGAGAG 681
 QY 181 AsnSerSerLeuAspGlnMetPheArgAlaLeuPheAlaMetAlaLysLeuProSerGlu 200
 DB 682 AACGACGCTTGGACCGAGATGTTCCGTGCGCTCTTCCATGGCCAGAGCGCTAGGAG 741
 QY 201 MetSerProAspLeuHisArgLysValSerValGlnTyrCysAspValIleHisLysLys 220
 DB 742 ATGAGCCCGGCTGACCGCAAGGATGATGTCGAGTACGCGAGTACGCAACAGAGAG 801
 QY 221 AlaLeuArgAsnLysLysLeuLeuArgAlaGlySerGlyGlyGlyLysAspProGly 240
 DB 802 GCTGTGAGGAACAAGAGCTTCTGCGGCGGAGC---GGAAGCGGGGCGAGCCACAG 858
 QY 241 AspAlaPheGlyIleValAlaProPheAlaArgArgProSerValHisSerAspLeuMet 260
 DB 859 GATGCCCTTGGCATCTTGGGCGCCCTTGGTGGAGAGCCAGCGGCAACGAGCTTATG 918
 QY 261 TyrIleArgGluLysAlaSerAlaGlySerGlnAlaLysAspLysGluValIle 280
 DB 919 TACATTCGTGAAAAAACCACTGTGCGGAGCGAGCTAAGAGCAAGAGCGCCCTGTGTCATC 978

QY 281 Ser 281
 DB 979 AGT 981

RESULT 9

AF239157 1616 bp mRNA linear ROD 06-JUL-2001
 LOCUS Rattus norvegicus DEXRAS1 (Dextral) mRNA, complete cds.
 DEFINITION AF239157
 ACCESSION AF239157
 VERSION AF239157.1 GI:7230767
 KEYWORDS
 SOURCE Rattus norvegicus.
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE

1 (bases 1 to 1616)
 Fang, M., Jaffrey, S.R., Sawa, A., Ye, K., Luo, X. and Snyder, S.H.
 Dextral: a G protein specifically coupled to neuronal nitric oxide
 synthase via CAPON
 Neuron 28 (1), 183-193 (2000)
 JOURNAL MEDLINE 20537828
 PUBMED 11086993
 REFERENCE 2 (bases 1 to 1616)
 Fang, M., Jaffrey, S.R., Sawa, A., Ye, K. and Snyder, S.H.
 Direct Subunit
 Submitted (24-FEB-2000) Neuroscience, Johns Hopkins University, 725
 N. Wolfe St, Baltimore, MD 21205, USA
 JOURNAL Location/Qualifiers

FEATURES

source
 1..1616
 /organism="Rattus norvegicus"
 /db_xref="taxon:10116"
 1..1616
 /gene="Dextral"
 118..960
 /gene="Dextral"
 /note="Ras-related protein; small GTPase; similar to Mus
 musculus DEXRAS1 encoded by GenBank Accession Number
 AF009246"
 /codon_start=1
 /product="DEXRAS1"
 /protein_id="AA043090.1"
 /db_xref="GI:7230768"
 /translation="MKLAAMTKMKCPDSELSIPAKNKRYMVLIGSSVKGTAIVSR
 ITGREDAVTPTEDEHFRKFSYIRGEVYOLDITDSCNHPBPAMRLSITGDVILY
 FSLDNRDSFEVQRLKQOIIDTKSLNKRKENDVPLVTCGNKGDYFREVQREI
 EQLVDDPQRCATFEISAKRNSLDQRFALPSEMSPDILHRKVSQYCDVILH
 KRALNRKRLLRAGSGGGHDGDAFGIILAPARRPSVSHDLMYIREKTSVSOAKDKER
 CVIS"

BASE COUNT 403 a 436 c 413 g 363 t 1 others
 ORIGIN

Alignment Scores:

Pred. No.: 1.43e-122 Length: 1616
 Score: 1399.50 Matches: 273
 Percent Similarity: 98.22% Conservative: 3
 Best Local Similarity: 97.15% Mismatches: 4
 Query Match: 96.58% Indels: 1
 DB: 10 Gaps: 1

US-09-709-103-2 (1-281) x AF239157 (1-1616)

QY 1 MetLysLeuAlaAlaMetIleLysLysMetCysProSerAspSerGluLeuSerIlePro 20
 DB 118 ATGAAGTGGCCGGCATGATCAAGAAAGTGGCCCAACCGCATCTGAACTGATATCCG 177
 QY 21 AlaLysAsnCysTyrArgMetValIleLeuGlySerSerLysValGlyLysThrAlaIle 40
 DB 178 GCCAAGAACTGCTACAGATGATGATCTCCGCTCATCCAAAGTGGCAGAGCGCCATC 237
 QY 41 ValSerArgPheLeuThrGlyArgPheGluAspAlaTyrThrProThrIleGluAspPhe 60

```

Db 238 GTGTGGCGCTTCTCAGAGGGCGCTTCAAGAGCGCTTACACCCCTTACACCTTGAAGACTTC 297
Qy 61 HisArgLysPheTyrSerIleArgGlyGluValTyrGlnLeuAspIleLeuAspThrSer 80
Db 298 CACGGAAGAGTTTACCTCGATCCCGGCGAAGTCTTACAGTGTGACATCTGACACATCT 357
Qy 81 GlyAsnHisProPheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIle 100
Db 358 GGCACATCATCCGTTTCCCGGCGACCGCGCTCTCTATCTCAGAGAGACGCTTTTCATT 417
Qy 101 LeuValPheSerLeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGlnGln 120
Db 418 CTGTGTGCTGCTTACGACACCGCGACTCTTCTGAGAGAGTGCAAGGCTCAACAGAG 477
Qy 121 IleLeuAspThrLysSerCysLeuLysAsnLysThrLysGlnAsnValAspValProLeu 140
Db 478 ATCTTACACACCAAGTCTCTCTCAAGACAAACCAAGAGATGTGGACGTGCGCGCT 537
Qy 141 ValIleCysGlyAsnLysGlyAspArgAspPheTyrArgGluValAspGlnArgGluIle 160
Db 538 GTCATTTGCGGTAAACAAGGGGACCGGACTTCTACCGCGAAGTGAGAGCGAGATT 597
Qy 161 GluGlnLeuValGlyAspAspProGlnArgCysAlaTyrPheGluIleSerAlaLysLys 180
Db 598 GACACAGCTGTGGGCGGAGACCTCAGCGCTGTGCTTCTGAGATCTCGCCAAAGAG 657
Qy 181 AsnSerSerLeuAspGlnMetPheArgAlaLeuPheAlaMetAlaLysLeuProSerGlu 200
Db 658 AATAGAGCGCTTGACGAGATGTTCCGTGGCTTTCGCTTCCGCAAGCGCTTACGAG 717
Qy 201 MetSerProAspLeuHisArgLysValSerValGlnTyrCysAspValLeuHisLysLys 220
Db 718 ATGAGCCCTGACTTGCACCCCAAGGTCTCTGTCTGACAGTCTGACGCTGCAACAAAG 777
Qy 221 AlaLeuArgAsnLysLysLeuLeuArgAlaGlySerGlyGlyGlyGlyValAspProGly 240
Db 778 GCTTGTGAGACAAAGAGCTTCTGCGGGGAGC---GAGAGTGGGGGCGACACGGA 834
Qy 241 AspAlaPheGlyIleValAlaProPheAlaArgArgProSerValHisSerAspLeuMet 260
Db 835 GATGCTTTGGCATCTTGGGCGCTTTCGTCGACAGCTTACGCGTATGACGACCTCATG 894
Qy 261 TyrIleArgGlyLysAlaSerAlaGlySerGlnAlaLysAspLysGluArgCysValIle 280
Db 895 TACATTCGTGAGAAACAGAGTGCAGCAGCAGCTTAAGGACAAAGAGCGCTGTCTATC 954
Qy 281 Ser 281
Db 955 AGT 957

RESULT 10
AF222979 4990 bp DNA linear PRI 02-JAN-2001
LOCUS Homo sapiens activator of G-protein signaling gene, complete cds.
DEFINITION AF222979
ACCESSION AF222979.1 GI:12004991
VERSION AF222979.1
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 4990)
AUTHORS Cismowski,M.J., Xie,X. and Duzic,E.
TITLE Genomic sequence of the human ras-related G-protein activator AGS1
JOURNAL Unpublished
REFERENCE
2 (bases 1 to 4990)
AUTHORS Cismowski,M.J., Xie,X. and Duzic,E.
TITLE Direct Submission
JOURNAL Submitted (10-JAN-2000) OSI Pharmaceuticals, 777 Old Saw Mill River
Road, Tarrytown, NY 10591, USA
FEATURES
SOURCE 1..4990
/organism="Homo sapiens"

```

```

mRNA
CDS
/db_xref="taxon:9606"
/chromosome="17"
/map="17p11.2"
/clone="BamHI-EcoRI subclone of p1"
/note="chromosomal localization verified by fluorescence
in-situ hybridization analysis"
join(2762..3188,3400..4643)
/product="activator of G-protein signaling"
join(2903..3188,3400..3959)
/note="ras-related protein; AGS1"
/codon_start=1
/product="activator of G-protein signaling"
/protein_id="AAG44256.1"
/db_xref="GI:12004992"
/translation="MKLAAMIKKCPDSBELSIPKNCYRMVILGSSVGTATVSRF
LTGRFEDAVPTIEDPFRKFEYSIGREYVOLDIDTSGNHPFAMRLSILTDVFLV
PSLNRPSFEVORLROOILDTKSLCENEDVPLVCGKNGDRDFREVDREI
EQIYGDPPORCAVPEISKKNSISDQMFRAIFPMAXLPSMSPDILHKVSVQCDVH
KKALRNKRLNLAGSGGGGGGDEGDAFGVAFAPRRPVSVDLMTIREKASAGDAKDE
KCVIS"
BASE COUNT 1049 a 1507 c 1509 g 925 t
ORIGIN
Alignment Scores: 8.38e-121 Length: 4990
Pred. No.: 1387.00 Matches: 281
Score: 79.83% Conservative: 0
Percent Similarity: 79.83% Mismatches: 0
Best Local Similarity: 95.72% Indels: 71
Query Match: 9 Gaps: 1
US-09-709-103-2 (1-281) x AF222979 (1-4990)
Qy 1 MetLysLeuAlaAlaMetIleLysLysMetCysProSerAspSerGluLeuSerIlePro 20
Db 2903 ATGAAACTGGCGCATGTATCAAGAAAGATGTCCGAGAGATCGGAGTGTATCCG 2962
Qy 21 AlaLysAsnCysTyrArgMetValIleLeuGlySerSerLysValGlyLysThrAlaIle 40
Db 2963 GCCAAGACCTGATCGCATGTGTCTATCCCTGCTCCAGGTGGCGCAAGCGGCATC 3022
Qy 41 ValSerArgPheLeuThrGlyArgPheGluAspAlaTyrThrProThrIleGluAspHe 60
Db 3023 GTGTGGCTTCTCTCAACCGCGCTTCAAGAGCGCTTACAGCTTCAACATCGAGACTTC 3082
Qy 61 HisArgLysPheTyrSerIleArgGlyGluValTyrGlnLeuAspIleLeuAspThrSer 80
Db 3083 CACCGCAAGTTCTACTCCATCCGCGGCGAGGTCTACACCTCGACATCTCGACACGTCC 3142
Qy 81 GlyAsnHisProPheProAlaMetArgArgLeuSerIleLeuThr----- 95
Db 3143 GGCACACACCCGTTTCCCGCCATGCGCGCTCTCCATCTCAGAGTGGAGCGGGGCC 3202
Qy 95 ----- 95
Db 3203 GGGCAGGTGGGAGGAGGAGGAGGCGGGAACCTCGGCGAGGGCGCCGCGAGCGCGGT 3262
Qy 95 ----- 95
Db 3263 CGGGCTGCGCGCGCGGAGTGTGCGCTTACGAGAGGCTTAGCGCGCGCGCGCGG 3322
Qy 95 ----- 95
Db 3323 GCCTCAAGTACACCCGACTTGTCCCTGGCGCGCCACCTGACCTTTCCTGTGCT 3382
Qy 96 -----GlyAspValPheIleLeuValPheSerLeuAspAsnArgAspSe 110
Db 3383 CTCTGTGCCCCCTCTAGGAGACGTTTTCATCTGTTGTTCACTGTGACACCGGACTC 3442
Qy 110 rPheGluGluValGlnArgLeuArgGlnIleLeuAspThrLysSerCysLeuLysAs 130
Db 3443 CTTCGAGAGAGTGCAGCGGCTGAGCAGAGATCTCGACACCAAGTCTTGCTCAAGAA 3502

```


QY 170 GCysAlaTyrPheGluIleSerAlaLysAsnSerSerLeuAspGlnMetPheArgAl 190
 Db 1360 CTGGCGCTACTTTCAGATCTCGGCCAAGAAGACAGACCTCGACCAATGTTCCGGCC 1301
 QY 190 aleuphealmetalaLysLeuProSerGlnMetSerProAspLeuHisArgLysValSe 210
 Db 1300 GCTTTCGCCATGCGCAAGTGTGCCAGATGAGCCACGACCTTCACCGCAAGGTCTTC 1241
 QY 210 rValGlnTyrCysAspValLeuHisLysLysAlaLeuArgAsnLysLysLeuLeuArgAl 230
 Db 1240 GGGGCGAGTACTGCGAGCTGTCACAAAGAGCGCTGCGGAACAAAGACTGCTGCGGGC 1181
 QY 230 aglySerGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 250
 Db 1180 CGGACAGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGC 1121
 QY 250 aarGatgProSerValHisSerAspLeuMetTyrIleArgGluLysAlaSerAlaGlySe 270
 Db 1120 GCGCGCGCCAGCGTACACAGCAGCTCATGTACATCCGCGAAGGCCAGCGCGCGGAG 1061
 QY 270 rGlnAlaLysAspLysGluArgGlyValIleSer 281
 Db 1060 CCAGGCCAAGGACAGAGCGCTGCTCATACG 1027

RESULT 12
 AF262018 5141 bp DNA linear PRI 31-MAY-2000
 LOCUS AF262018
 DEFINITION Homo sapiens dexamethasone-induced ras-related protein 1 (DEXRAS1)
 ACCESSION AF262018
 VERSION AF262018.1 GI:8118456
 KEYWORDS gene, complete cds.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 5141)
 AUTHORS Kempainen, R.J. and Behrend, E.N.
 JOURNAL Human Dextral gene
 TITLE Unpublished
 REFERENCE 2 (bases 1 to 5141)
 AUTHORS Kempainen, R.J. and Behrend, E.N.
 JOURNAL Direct Submission
 TITLE Submitted (30-APR-2000) Anatomy & Physiology, Auburn University,
 JOURNAL College of Veterinary Medicine, Auburn, AL 36832, USA
 FEATURES
 source
 1..5141
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="2"
 /map="2q32"
 <2548..>3604
 /gene="DEXRAS1"
 /join(<2548..2833,3045..>3604)
 /gene="DEXRAS1"
 /product="dexamethasone-induced ras-related protein 1"
 /join(2548..2833,3045..3604)
 /gene="DEXRAS1"
 /note="Dextral"
 /codon_start=1
 /product="dexamethasone-induced ras-related protein 1"
 /protein_id="AAF72997.1"
 /db_xref="GI:8118456"
 /translation="MKLAAMIKKMPDSELSIPAKNCYRMVILGSSKVGKTAIVSRF
 ITCRPEADATPTTDFHRRKYSIRGEVYGLDIDNSGNHPPAMRLSTLIGDVFILY
 FSLDNDSFEVORLKOQLIDTRSCSKKNTKENVDPYIVICGNKGDRDYRDDREI
 EOLVGDPOKCAFEISAKNSLDQMFALPAMAKLPSEMSPDHKKVSVYCDVLH
 KKAIRNKKILIRAGSGGGGDPDAGFIVAFARRPVSHDLVYIRKASAGSOADKE
 RCVIS"
 BASE COUNT 1090 a 1579 c 1522 g 950 t
 ORIGIN

Alignment Scores:
 Pred. No.: 1,67e-120 Length: 5141
 Score: 1384.00 Matches: 280
 Percent Similarity: 79.83% Conservative: 1
 Best Local Similarity: 79.55% Mismatches: 0
 Query Match: 95.51% Indels: 71
 DB: 9 Gaps: 1

US-09-709-103-2 (1-281) x AF262018 (1-5141)

QY 1 MetLysLeuAlaAlaMetIleLysLysMetCysProSerAspSerGlnLeuSerIlePro 20
 Db 2548 ATGAAGACTGGCGCGATGATGAAGAAAGATGTGCCGAGCGACTGAGATGATCCG 2607
 QY 21 AlalysAsnCysTyrArgMetValIleLeuGlySerSerLysValGlyLysThralaIle 40
 Db 2608 GCCAAGACTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2667
 QY 41 ValSerArgPheLeuThrGlyArgPheGluAspAlaTyrThrProThrIleGluAspPhe 60
 Db 2668 GTGTGCGCTTCCTTCACCGCGCGCTTCGAGGACCCCTACAGCCTACCATGAGACTTC 2727
 QY 61 HisArgLysPheTyrSerIleArgGlyGluValTyrGlnLeuAspIleLeuAspThrSer 80
 Db 2728 CACCGAAGTCTTACTCATCTCATCTCATCTCATCTCATCTCATCTCATCTCATCTCAT 2787
 QY 81 GlyAsnHisProPheProAlaMetArgArgLeuSerIleLeuThr 95
 Db 2788 GGCACACACCGCTTCCCGGCATCGCGGCGCTTCACATCTTCACAGAGTGACCGGGGCC 2847
 QY 95 2848 GGCAGCTCCCGAGAGGAGGCGGGGAAACCTCGCGAGCGCCCGCGAGCGCGGT 2907
 QY 95 2908 CCGGCTGCCCGCGCGCGAGTACTGCGCTTCGAGAGGCTAGGCGCCCGCGCGG 2967
 QY 95 2968 GCGTCAAGTACGCCGACTTGTCCCTGGCGCGCACCCCTTCCTCTTCTCTCT 3027
 QY 96 3028 CTCTGTGCCCTCTAGAGAGAGCTTTTCATCTGTGTTCAGTGTGCAACCGCGACTC 3087
 QY 110 rPheGluGluValGlnArgLeuArgGlnGlnIleLeuAspThrLysSerCysLeuLysAs 130
 Db 3088 CTTCAGAGAGGTGACAGCGGCTCAACAGACAGATCTCGACACCAAGTCTTGCCTCAAGA 3147
 QY 130 nLysThrLysGluAsnValAspValProLeuValIleCysGlyAsnLysGlyAspArgAs 150
 Db 3148 CAAACCAAGGAGAGAGAGTGTGAGCTGCTGTCATCTGCGGCAACAGGGGTGACCGCGA 3207
 QY 150 pPheTyrArgGluValAspGlnArgGlnIleGlnIleLeuValGlyAspAspProGlnArg 170
 Db 3208 CTTCACCGCGAGGTGAGACCGAGAGTGTGAGAGCTGGGGCGAGCGACCGCCGCG 3267
 QY 170 gCysAlaTyrPheGluIleSerAlaLysLysAsnSerSerLeuAspGlnMetPheArgAl 190
 Db 3268 CTGGCGCTACTTTCAGATCTCGGCCAAGAAGACAGACCTCGACCAAGTTCGCGGC 3327
 QY 190 aleuphealmetalaLysLeuProSerGlnMetSerProAspLeuHisArgLysValSe 210
 Db 3328 GCTTTCGCCATGCGCAAGTGTGCCAGATGAGCCACGACCTTCACCGCAAGGTCTTC 3387
 QY 210 rValGlnTyrCysAspValLeuHisLysLysAlaLeuArgAsnLysLysLeuLeuArgAl 230
 Db 3388 GGGGCGAGTACTGCGAGCTGTCACAAAGAGCGCTGCGGAACAAAGACTGCTGCGGGC 3447
 QY 230 aglySerGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 250
 Db 3448 CGGACAGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGC 3507

QY 250 aargargproserValHisSeraspLeumethyrrileargGluLysAlaSerAlaGlyse 270
 DB 3508 GGGCCGGCCACAGGACAGGACCTCATGATCATCCGAGAGAGCCAGCGCGGAG 3567

QY 270 rGlnAlaLysAspLysGluArgCysValIleSer 281
 DB 3568 CCAGGCCAAGACAGGAGCGCTGCTCATCAGC 3601

RESULT 13
 AL603710/c
 LOCUS AL603710 179124 bp DNA linear ROD 05-APR-2002
 DEFINITION Mouse DNA sequence from clone Rp23-247B13 on chromosome 11,
 complete sequence.
 ACCESSION AL603710
 VERSION AL603710.8 GI:20068514
 KEYWORDS HTG.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
 1 Almeida, J.
 TITLE Direct Submission
 JOURNAL Submitted (04-APR-2002) Wellcome Trust Sanger Institute, Hinxton,
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
 humquerry@sanger.ac.uk Clone requests: clonequest@sanger.ac.uk
 On Apr 7, 2002 this sequence version replaced gi:18070899.
 During sequence assembly data is compared from overlapping clones.
 Where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the
 variation annotation may not be found in the sequence submission
 corresponding to the overlapping clone, as we submit sequences with
 only a small overlap as described above.
 This sequence was either double-stranded or sequenced with an alternate
 chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by at least
 one plasmid subclone or more than one M13 subclone; and the
 assembly was confirmed by restriction digest. The following
 abbreviations are used to associate primary accession numbers given
 in the feature table with their source databases: Em, EMBL; Sw, SWISSPROT; Tr, TrEMBL; Wp, WormPeP; Information on the WormPeP
 database can be found at
 http://www.sanger.ac.uk/Projects/C_elegans/wormpep/Rp23-247B13.html
 from the RPI-23 Mouse PAC Library
 constructed by the group of Pieter de Jong.
 For further details see http://www.chori.org/bacpac/home.htm
 VECTOR: pBAC6.6.

FEATURES
 source Location/Qualifiers
 1. 179124
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /chromosome="11"
 /clone="Rp23-247B13"
 /clone_11b="RPI-23"
 BASE COUNT 46162 a 43228 c 43736 g 45998 t
 ORIGIN

Alignment Scores:
 Pred. No.: 2,48e-117 Length: 179124
 Score: 1370.00 Matches: 274
 Percent Similarity: 83.43% Conservatave: 3
 Best Local Similarity: 82.53% Mismatches: 3
 Query Match: 94.55% Indels: 52
 DB: 10 Gaps: 2

US-09-709-103-2 (1-281) x AL603710 (1-179124)

QY 1 MetLysLeuAlaAlaMetIleLysMetCysProSerAspSerGluLeuSerIlePro 20
 DB 122425 ATGAAACTGGCCGCGATGATCAAGAAAGATGTCCCAAGGAGCTGTGAAGTATCCCG 122366

QY 21 AlAlaLysAnCysTyrArgMetValIleLeuGlySerSerLysValGlyLysThrAlaIle 40
 DB 122365 GCCAAGAACTGCTACAGGATGGTATCTCGGCTCATCAAGAGGGGCAAGAGCCGCAATT 122306

QY 41 ValSerArgPheIleuThrGlyArgPheGluAspLysPheIleuThrProThrIleGluAspPhe 60
 DB 122305 GTGTGCGCTTCTCAGCGGGCGGTTTCAGAGATGCTTACACCCCTACATCAGAGGACTTC 122246

QY 61 HisArgLysPheTyrSerIleArgGlyValIleTyrGlnLeuAspIleLeuAspThrSer 80
 DB 122245 CACCGAAGTTTACTGATCATCCGCGGCGCAAGTCTACAGTTGACATACATCAGATCC 122186

QY 81 GlyAsnHisProPheProAlaMetArgIleuSerIleLeuThr 95
 DB 122185 GGCATCATCCTGTTCCCGCATGGGCGCTCTATCTCAGAGGAGTGGAGATC 122126

QY 95 95
 DB 122125 GACGGGATPAGGGGAGCCGAGTCTTGAGGGAGTGATGGGCTGTGTGTGCTTG 122066

QY 95 95
 DB 122065 GGGCTTGGCCCTTACCTGGGGGTGTCTGACCTGTGTGCGGAGCTGTCTACCTTT 122006

QY 96 96
 DB 122005 CCCACTGCTTCCCTTATAGGAGAGCTTTTCAATCTGGGTGCTACGTTAGACACCGGAGC 121946

QY 110 SerPheGluGluValGlnArgLeuArgGlnGlnIleLeuAspThrLysSerCysLeuLys 129
 DB 121945 TCATTTCGAGAGGTCGCAAGGCTCAAAACAGATCTCAGACACCAAGTCTGTCTCAG 121886

QY 130 AsnLysThrLysGluAsnValAspValProLeuValIleCysGlyAsnLysGlyAspArg 149
 DB 121885 AACAAACCAAGAAAGAAATGTCGACCTGCGCTGTCTATTCGTTAACAAAGGGAGCCGG 121826

QY 150 AspPheTyrArgGluValAspGlnArgIleLeuGlnLeuValGlyAspAspProGln 169
 DB 121825 GACTTCTACCGGGAAGTAGACAGCGGAGATGAGAGCTGTGGGTGAGACCTCTCAG 121766

QY 170 ArgCysAlaTyrPheGluIleLeuSerAlaLysLysAsnSerSerLeuAspGlnMetPheArg 189
 DB 121765 CGTTGTGCTCTCTTCGATTCGATTCAGCCAAAGAAACAGCACTTGACACATCTTCCTGT 121706

QY 190 AlaLeuPheAlaMetAlaLysLeuProSerGlnuSerProAspLeuHisArgLysVal 209
 DB 121705 GCGCTTGTGCTTCATGCGCAAGCTGCTTAGCGAGATGAGACCCCGCACTTGACCGAAGGTA 121646

QY 210 SerValGlnTyrCysAspValIleuHisLysLysAlaLeuArgAsnLysLysLeuLeuArg 229
 DB 121645 TCTGTGCGAGTACTGCGAGCTACTGCAACAAAGAGCTGTGAGAGCAAGAAAGCTTCTCGT 121586

QY 230 AlaGlySerGlyGlyGlyGlyGlyAspProGlyLysPheGlyIleValAlaIleProPhe 249
 DB 121585 GCGGGGACAC---GGAGGGGGGGGGGAGCACCGCATGCTTTGGCAATCTGGGCCCTTT 121529

QY 250 AlaArgArgProSerValHisSerAspLeuMetTyrIleArgGluLysAlaSerAlaGly 269
 DB 121528 GCTCGCAGACCCAGCGCTGCAACGCGACCTCATGTACATTCCTGAAAAAACCAAGTGTGCGC 121469

QY 270 SerGlnAlaLysAspLysGluArgCysValIleSer 281
 DB 121468 AGCCAGGCTAAGAGCAAGAGAGCGCTGTCTCATAGT 121433

RESULT 14
 LOCUS AC025909/c 247899 bp DNA linear HTG 27-JUN-2001
 DEFINITION Mus musculus chromosome 11 clone Rp23-82E8, WORKING DRAFT SEQUENCE,
 6 unordered pieces.
 ACCESSION AC025909
 VERSION AC025909.28 GI:14547761
 KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.

Db 19030 GCGGCGAGC---GGAGCGGGGGGCGACACCGGCGATGCGCTTGCGATCTTGCGGCGCTTT 18974
 QY 250 AlAaGArGProSeRvAlHISserAPLeMeTtYrIlleRgSlUySAlaSeRaLaGly 269
 Db 18973 GCTCGGAGACCGCGCTGCGACGACCTCATGTCATTCGTCGTAACCAAGCTGCGGC 18914
 QY 270 SerGlnAlaLysApLySgIuArGcYsValIlleSer 281
 Db 18913 AGCGACGCTTAGCAACAGAGCGCTGTGCATCAGT 18878
 RESULT 15
 AC090608 183598 bp DNA linear HTG 22-MAR-2002
 LOCUS Homo sapiens chromosome 17 clone RP11-524F11 map 17, *** SEQUENCING
 DEFINITION IN PROGRESS ***, 2 unordered pieces.
 AC090608 183598 bp DNA linear HTG 22-MAR-2002
 AC090608 183598 bp DNA linear HTG 22-MAR-2002
 HTG: HTGS_PHASE1, HTGS_FULLTOP, HTGS_ACTIVERFN.
 KEYWORDS Homo sapiens.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 183598)
 Birren, B., Linton, L., Nusbaum, C. and Lander, E.
 Homo sapiens chromosome 17, clone RP11-524F11
 Unpublished
 2 (bases 1 to 183598)
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, S.,
 Barina, N., Bastien, V., Boguslavsky, L., Boukhalter, B., Brown, A.,
 Camarata, J., Campopiano, A., Choepel, J., Colangelo, M., Collins, S.,
 Collymore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S.,
 Dodge, S., Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J.,
 Gardyna, S., Ginde, S., Goyette, M., Graham, W., Grand-Pierre, N.,
 Hages, B., Heatford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,
 Jones, C., Karatas, A., Laroque, K., Lamazares, R., Landers, T.,
 Lehoczy, J., Levine, R., Liu, G., Maclean, C., Macdonald, P.,
 Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K.,
 McPheeters, R., Meldrum, J., Menes, L., Mihova, T., Mlenga, V.,
 Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C.H.,
 O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,
 Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R.,
 Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M.,
 Roy, A., Santos, R., Schauer, S., Schupack, R., Seaman, S., Severy, P.,
 Sounez, C., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
 Strauss, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J.,
 Travers, M., Travis, N., Triglio, J., Vassiliev, H., Viel, R., Vo, A.,
 Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J.,
 Zembek, L., Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (03-MAR-2001) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 183598)
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N.,
 Anderson, S., Barina, N., Bastien, V., Bloom, T., Boguslavsky, L.,
 Boukhalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J.,
 Chazaro, B., Choepel, J., Colangelo, M., Collins, S., Collymore, A.,
 Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
 Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,
 Ginde, S., Goid, S., Goyette, M., Graham, W., Grand-Pierre, N.,
 Hages, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
 Kamet, A., Karatas, A., Kells, C., Laroque, K., Lamazares, R.,
 Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G.,
 Maclean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,
 McEwan, P., McKernan, K., Meldrum, J., Menes, L.,
 Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicot, R.,
 Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,
 Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
 Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
 Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupack, R.,
 Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
 Strauss, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J.,
 Travers, M., Travis, N., Triglio, J., Vassiliev, H., Viel, R., Vo, A.,
 Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J.,
 Zembek, L., Zimmer, A. and Zody, M.

TITLE
 JOURNAL
 COMMENT
 Viell, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,
 Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (22-MAR-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Mar 22, 2002 this sequence version replaced gi:16445176.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html
 ----- genome center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WITR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L11985
 Center clone name: 524_F_11

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 2 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 * 1 178215: contig of 178215 bp in length
 * 178216 178315: gap of 100 bp
 * 178316 183598: contig of 5283 bp in length.
 Location/Qualifiers
 1. 183598
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="17"
 /map="17"
 /clone="RP11-524F11"
 /clone_lib="RPCT-11 Human Male BAC"
 BASE COUNT 40950 a 47626 c 49185 g 45365 t 272 others
 ORIGIN
 Alignment Scores:
 Pred. No.: 5,34e-116 Length: 183598
 Score: 1356.00 Matches: 279
 Percent Similarity: 78.81% Conservative: 0
 Best Local Similarity: 78.81% Mismatches: 2
 Query Match: 93.58% Indels: 73
 DB: 2 Gaps: 1
 US-09-709-103-2 (1-281) x AC090608 (1-183598)
 QY 1 MetLysLeuAlaAlaMetIleLysLysMetCysProSerAspSerGluLeuSerIlePro 20
 Db 181508 ATGAAACTGGCGCGGATGATCAAGAAATGTCGCCGAGCGACTGGAGCTAGATATCCG 181567
 QY 21 AlAluSaNCySTYrArGMeTValIlleLyuLysSerLysValGlyLysThAlaIlle 40
 Db 181568 GCCAAGAACTGATCGATGATGATCTCGGCTCGTCCAGAGTGCGCAACAGCCGATC 181627
 QY 41 ValSerArGpHeLeThGlyArGpHeGluAspAlaTyTrProThrIleGluAspHe 60
 Db 181628 GTGTGCCCTTCCTCACCAGCGGCTTCGAGAGAGCCCTTACACCTGAGAGACTTC 181687
 QY 61 HisArGlySpHeTySerIleArGlyGlyValTyrgLnuAspIleLeuAspThrSer 80
 Db 181688 CACCCAGATCTTACTCATCTACCTCGGCGAGAGTGTACAGCTGCAGATCTCGACACGTCG 181747
 QY 81 GlyAsnHISpHeProAlaMetArGlyLeuSerIleLeuThr 95
 Db 181748 GGCACACACCGCTTCCCGCATGGGCGCTTCATCTCACAGAGTGAGCGGGGCC 181807
 QY 95 ----- 95
 Db 181808 GGGCAGTGGCGGAGGAGGCGGGAACCTTGGCCAGGCGCCCGCGAGCGCGGCT 181867

QY 95 ----- 95
Db 181868 CCGGCTGCCGGCGCCGAGTAGTGCGCTTGCGCTTAGAGAGGCTAGCGCGCCCGCGCG 181927
QY 95 ----- 95
Db 181928 GCCTCAAGTCAGCCCGACTGTGCCCTGGGGCGGCCACCTCACCCTTCTCTTCTGCT 181987
QY 96 -----GlyAspValPheIleuValPheSerLeuAspAsnArgAspSe 110
Db 181988 CTCTGTGCCCTCTAGGAGACGTTTTCATCTGCTGTTCAGTCTGGACACACCGGACTTC 182047
QY 110 rPheGluGluValGlnArgLeuArgGlnGlnIleLeuAspThrIysSerCysLeuLysAs 130
Db 182048 CTTGAGGAGGTGCGCGGCTAGGACAGACAGATCTCGACACCAAGTCTTGCTCAAGAA 182107
QY 130 nLysThrIysGlu--AsnValAspValProLeuValIleCysGlyAsnLysGlyAspArg 149
Db 182108 CAAAACCAAGGAGANACTNGSAGCTGGCCCTGGTCACTCTGCGCAACAGGTGACCGC 182167
QY 150 AspPheTyArgGluValAspGlnArgGluIleGluGlnLeuValGlyAspAspProGln 169
Db 182168 GACTTCTACCGGAGGTGAGCGGACGAGATCGAGCTGTGTGGCGACGACCCCGAG 182227
QY 170 ArgCysAlaItyrPheGluIleSerAlaLysAsnSerSerLeuAspGlnMetPheArg 189
Db 182228 CGCTGGCTTACTTCTGAGATCTCGGCAAGAAAGAACAGACCTGGACAGATGTTCGCG 182287
QY 190 AlaLeuPheAlaMetAlaLysLeuProSerGluMetSerProAspLeuHisArgLysVal 209
Db 182288 GGGCTCTTGGCCATGGCCAGCTGGCCAGGAGATGAGCCAGACCTGCACGCGCAGSTC 182347
QY 210 SerValGlnItyrCysAspValLeuHisLysLysAlaLeuArgAsnLysLysLeuLeuArg 229
Db 182348 TCGGTGCAGTACTGCGACGTGCTGCACAAAGGCGCTCGGAACAAAGACCTGCTGCGG 182407
QY 230 AlaGlySerGlyGlyGlyGlyGlyAspProGlyAspAlaPheGlyIleValAlaProPhe 249
Db 182408 GCGGCGAGCGGGCGGGCGGGCGGAGCCCGGCGCTTGGCATGTGGCACCTTTC 182467
QY 250 AlaArgArgProSerValHisSerAspLeuMetTyrIleArgGluLysAlaSerAlaGly 269
Db 182468 GCGGCGCGGGCCAGCGTACACAGGACCTCATGTACATCCGCGAAGAGCCAGCGCGCGC 182527
QY 270 SerGlnAlaLysAspLysGluArgCysValIleSer 281
Db 182528 AGCCAGGCCAAAGACAGAGCGCTGCTCATCAGC 182563

Search completed: December 28, 2002, 02:13:33
Job time : 3319 secs

. . .

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: December 30, 2002, 14:39:31; Search time 353 Seconds

(without alignments)
1799.047 Million cell updates/sec

Title: US-09-709-103-1f1

Perfect score: 1450
Sequence: 1 MKLAMIWKMCPSDSELSIP.....REKASAGSOAKXERCVIS. 282

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODL=frame+ p2n.model -DEV=xlp
-O=/cgn2.1/USPTO_TRANSDS0709103/runat.30122002.143925.20642/app.query.fasta.1.455
-DB=N_Geneseq.101002 -QFMT=fastad -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPTCT=0
-DOOEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blom62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pcio -NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=TRANSUS0909103 @cgn 1.1.79 @runat.30122002.143925.20642 -NCPU=6 -ICPU=3
-NO_XLPRY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELop=6 -DELEXt=7

Database :

N_Geneseq.101002:*
1: /SID32/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SID32/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SID32/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
4: /SID32/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SID32/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SID32/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SID32/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SID32/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SID32/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SID32/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SID32/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SID32/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SID32/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
14: /SID32/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
15: /SID32/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
16: /SID32/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
17: /SID32/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
18: /SID32/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
19: /SID32/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
20: /SID32/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
21: /SID32/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
22: /SID32/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
23: /SID32/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
24: /SID32/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1449	99.9	846	21	AAZ36892
2	1449	99.9	1740	21	AAZ36914
3	1449	99.9	1801	21	AAZ36893
4	1449	99.9	1841	20	AAZ23022
5	1441	99.4	1776	21	AAA49177
6	1399.5	96.5	1689	20	AAZ23024
7	1387	95.7	3986	20	AAZ23023
8	1343	92.6	3079	20	AAZ23025
9	893	61.6	837	21	AAZ36913
10	893	61.6	2832	24	ABL92076
11	893	61.6	2973	24	ABL92087
12	893	61.6	3427	24	ABK71563
13	867	59.8	3020	24	ABU92134
14	717	49.4	951	23	AA590571
15	615	42.4	1305	23	ABU07789
16	494	34.1	624	24	ABO28318
17	494	34.1	624	24	ABO28319
18	459	31.7	624	24	ABO28320
19	459	31.7	624	24	ABO28321
20	370	25.5	368	22	ABA51382
21	370	25.5	368	22	ABA69388
22	370	25.5	368	22	ABA36323
23	370	25.5	368	22	AAK17661
24	370	25.5	368	22	AAK43477
25	370	25.5	368	22	AA124259
26	370	25.5	368	22	AA149541
27	370	25.5	368	22	AA109818
28	370	25.5	368	24	AB517600
29	359.5	24.8	4543	23	ABU07788
30	334	23.0	1087	22	AA160838
31	334	23.0	1108	22	AA159052
32	334	23.0	4167	24	AAAD37605
33	330.5	22.8	960	23	ABU25241
34	332	22.2	3061	24	ABU27648
35	307.5	21.2	702	23	ABU14159
36	305	21.0	1144	22	AAH90075
37	305	21.0	4933	22	AAK81235
38	305	21.0	4934	22	AAK81234
39	304	21.0	1249	24	ABO72525
40	302	20.8	558	11	AAO03212
41	302	20.8	3300	22	AA158569
42	301	20.8	2205	24	ABU50199
43	301	20.8	2205	24	ABU50231
44	301	20.8	2277	24	ABU50202
45	301	20.8	2277	24	ABU50234

ALIGNMENTS

RESULT 1
AAZ36892
ID AAZ36892 standard; cDNA, 846 BP.

AAZ36892:
DT 13-MAR-2000 (first entry)

DE cDNA encoding an activator of G protein signalling (AGS) protein.

XX Activator of G protein signalling; AGS; ras-related G protein;

KW GTP hydrolysis; G protein activity; pheromone response pathway;

KM G protein-coupled signal transduction; G-gamma selectivity;

OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 1..846
 FT /*tag= a
 FT /product= "activator of G protein signalling (AGS)
 FT protein"
 XX
 PN WO9958670-A1.
 XX 18-NOV-1999.
 XX
 XX 07-MAY-1999; 99WO-US10151.
 PF 08-MAY-1998; 98US-0084842.
 PR 07-OCT-1998; 98US-0103355.
 XX
 XX (CADU-) CADUS PHARM CORP.
 PA
 XX Cismowski M, Duzic E;
 XX WPI; 2000-072337/06.
 DR P-PSDB; AAY53921.
 XX
 XX A new activator of G protein signalling used to treat disorders
 PT characterized by an aberrant AGS protein activity -
 XX
 PS Claim 3; Fig 3A; 162pp; English.
 XX
 CC The present sequence encodes an activator of G protein signalling (AGS)
 CC protein. The cDNA sequence was isolated from a human liver cDNA
 CC library. The AGS protein exhibits homology to ras-related G proteins,
 CC and contains alterations in conserved amino acids consistent with a
 CC deficiency in GTP hydrolysis activity. AGS stimulates G protein
 CC activity, G protein-coupled signal transduction and the pteronome
 CC response pathway in a receptor-independent manner. The AGS protein
 CC also shows G-gamma selectivity, as measured by growth assays in
 CC yeast expressing various mammalian G-gamma constructs, and
 CC tissue-specific expression, as measured by Northern blot analysis.
 CC The AGS protein can be used to screen for compounds that modulate
 CC cellular signal transduction. The protein is used to treat disorders
 CC characterized by an aberrant AGS protein activity or AGS nucleic acid
 CC expression.
 XX
 SQ Sequence 846 BP; 176 A; 288 C; 252 G; 130 T; 0 other;

Alignment Scores:
 Pred. No.: 6,91e-158 Length: 846
 Score: 1449.00 Matches: 281
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 99.93% Indels: 0
 DB: 21 Gaps: 0

US-09-709-103-1f1 (1-282) x AAZ236992 (1-846)
 Qy 1 MetLysLeuAlaMetIleLysLysMetCysProSerAspSerGluLeuSerIlePro 20
 Db 1 ATGAACCTGGCGGATGATCAAGAGATGTCCCGAGCGACTCGAGCTGAGTATCCCG 60
 Qy 21 AlalysAsnCysTyrArgMetValIleLeuGlySerSerLysValGlyLysThrAlaIle 40
 Db 61 GCCAAGAACTGCTATCGCATGTCTCGCTCGTCCAAAGGTGGCAAGACGGCCATC 120
 Qy 41 ValSerArgPheLeuThrGlyArgPheGluAspAlaTyrThrProThrIleGluAspPhe 60
 Db 121 GTGTGGCGCTTCTCCTACCGGCGCTTCAGAGACGCTACACGCTACCATCGAGGACTTC 180
 Qy 61 HisArgLysPheTyrSerIleArgGlyGluValTyrGlnLeuAspIleLeuAspThrSer 80
 Db 181 CACGCAAGTTCTACTCCATCCCGGGGAGGTCTACAGCTCGACATCTTCGACAGTCC 240
 Qy 81 GlyAsnHisProPheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIle 100
 Db 241 GGCAACCAACCGTTCCCGCCCATCGCGGCTCTCCATCTCTCACAGGAGAGCGTTTTCATC 300

Qy 101 LeuValPheSerLeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGlnGln 120
 Db 301 CTTGGTGTTCAGTCTCGACAACCGGACTCTTCGAGGAGGTGTCAGCGCTCAGGCGAG 360
 Qy 121 IleLeuAspThrLysSerCysLeuLysAsnLysThrLysGluAsnValAspValProLeu 140
 Db 361 ATCTCTCGACACCAAGTCTTGCTCTCAAGAACAAACCAAGGAGAACGTGACGTGCCCTG 420
 Qy 141 ValIleCysGlyAsnLysGlyAspArgAspPheTyrArgGluValAspGlnArgGluIle 160
 Db 421 GTCATCTGCGGCAACAAGGCTGACCGGACTTCTACCGCGAGGTGACCGCGAGATC 480
 Qy 161 GluGlnLeuValGlyAspAspProGlnArgCysAlaTyrPheGluIleSerAlaLysLys 180
 Db 481 GAGCAGCTGTGTGGCGAGACCCCCAGCGCTGCGCTACTTCGAGATCTCGGCCAAGAAG 540
 Qy 181 AsnSerSerLeuAspGlnMetPheArgAlaLeuPheAlaMetAlaLysLeuProSerGlu 200
 Db 541 AACAGCAGCTGGACCAAGATGTTCCGCGGCTCTTCGCAATGGCCAAAGCTGCCAGCGAG 600
 Qy 201 MetSerProAspLeuHisArgLysValSerValGlnTyrCysAspValLeuHisLysLys 220
 Db 601 ATGAGCCCAAGACTGACCGCAAGGTCTCGGTGCAGTACTGCGACGTGTCACCAAGAAG 660
 Qy 221 AlaLeuArgAsnLysLysLeuArgAlaGlySerGlyGlyGlyGlyGlyAspProGly 240
 Db 661 GCGCTGCGGAAACAAGAGCTGCTGCGGCGGCGAGCGCGGCGGCGGCGGCGGCGGCG 720
 Qy 241 AspAlaPheGlyIleValAlaProPheAlaArgArgProSerValHisSerAspLeuMet 260
 Db 721 GACGCTTTGGCATCTGTGGCACCCTTCGCGCGCGGCGGCGGCGGCGGCGGCGGCGG 780
 Qy 261 TyrIleArgGluLysAlaSerAlaGlySerGlnAlaLysAspLysGluArgCysValIle 280
 Db 781 TACATCCGCGAAGGCGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 840
 Qy 281 Ser 281
 Db 841 AGC 843

RESULT 2
 AAZ36914
 ID AAZ36914 standard; DNA; 1740 BP.
 XX AC AAZ36914;
 XX DT 13-MAR-2000 (first entry)
 XX DE DNA encoding an activator of G protein signalling (AGS) protein.
 XX KW Activator of G protein signalling; AGS; ras-related G protein;
 KW GTP hydrolysis; G protein activity; pheromone response pathway;
 KW G protein-coupled signal transduction; G-gamma selectivity;
 XX cellular signal transduction; ss.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 145..991
 FT /*tag= a
 FT /product= "activator of G protein signalling (AGS)
 FT protein"
 XX
 PN WO9958670-A1.
 XX 18-NOV-1999.
 XX 07-MAY-1999; 99WO-US10151.
 XX 08-MAY-1998; 98US-0084842.
 XX 07-OCT-1998; 98US-0103355.
 XX (CADU-) CADUS PHARM CORP.

XX Cismowski M, Duzic E;
XX MPI: 2000-072337/06.
DR P-PSDB; AAY53924.

XX A new activator of G protein signalling used to treat disorders
PT characterized by an aberrant AGS protein activity -
XX

PS Disclosure; Page 146-148; 162pp; English.

XX The present sequence encodes an activator of G protein signalling (AGS)
CC protein. The cDNA sequence was isolated from a human liver cDNA
CC library. The AGS protein exhibits homology to ras-related G proteins,
CC and contains alterations in conserved amino acids consistent with a
CC deficiency in GTP hydrolysis activity. AGS stimulates G protein
CC activity, G protein-coupled signal transduction and the pheromone
CC response pathway in a receptor-independent manner. The AGS protein
CC also shows G-gamma selectivity, as measured by growth assays in
CC yeast expressing various mammalian G-gamma constructs, and
CC tissue-specific expression, as measured by Northern blot analysis.
CC The AGS protein can be used to screen for compounds that modulate
CC cellular signal transduction. The protein is used to treat disorders
CC characterized by an aberrant AGS protein activity or AGS nucleic acid
CC expression.

SQ Sequence 1740 BP; 422 A; 546 C; 483 G; 289 T; 0 other;

Alignment Scores:

Pred. No.: 1,93e-157 Length: 1740
Score: 1449.00 Matches: 281
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.93% Indels: 0
DB: 21 Gaps: 0

US-09-709-103-1f1 (1-282) x AA236914 (1-1740)

QY 1 MetLysLeuAlaIaMetLysMetCysProSerAspSerGluLeuSerIlePro 20
DB 146 ATGAACTGGCCCGCATGATCAAGAGATGTGCCGAGCACTCGAGCTGAGTATCCCG 205
QY 21 ALlysAsnCySTyRArgMetValIleLeuGlySerSerLysValGlyThrAlaIle 40
DB 206 GCCAAGACGCTATGTCATGTCATCTCGGCTCGCCAGGTGGGCAAGCGCCATC 265
QY 41 ValSerArgPheLeuThrGlyArgPheGluAspAlaTyThrProThrIleGluAspPhe 60
DB 266 GTGTGGGCTTCTCTCAACCGCGCTTCGAGAGCGCTACACGCTTACCTGAGGACTTC 325
QY 61 HisArgLysPheTySerIleArgGlyGluValTyrgLnuAspIleLeuAspThrSer 80
DB 326 CACCGCAAGTTCATCTCATCCGGGCGAGGCTTACCGACTTCGACATCTTCGACACGTC 385
QY 81 GlyAsnHisProPheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIle 100
DB 386 GGCACCAACCGCTTCCCGCATGCGCGCTCTCCATCTCTCAAGAGACGTTTCATC 445
QY 101 LeuValPheSerLeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGlnGln 120
DB 446 CTGGTGTTCAGTCTGCAACCGGACTCTCTTCGAGAGAGGTGCAAGCGGCTCAGGCAAG 505
QY 121 IleLeuAspThrLysSerCysLeuLysAsnLysThrLysGluAsnValAspValProLeu 140
DB 506 ATCTTCAGACCAAGTCTTCTTCAGACAAACCAAGAGAACTGGACGTGCCCTG 565
QY 141 ValIleCysGlyAsnLysGlyAspArgAspPheTyArgLysValAspGlnArgGlnIle 160
DB 566 GTATCTCGCGCAACCAAGGTGACCGGACTTACCGGAGGTGACCGCGAGATC 625
QY 161 GluGlnLeuValGlyAspAspProGlnArgCysAlaTyThrGluIleSerAlaLysLys 180
DB 626 GAGCAGCTGTGGGCGAGACCGCCAGCGCTGAGATCTTGAGATCTTGCGCAAGAG 685

QY 181 AsnSerLeuAspGlnMetPheArgAlaLeuPheAlaMetAlaLysLeuProSerGlu 200
DB 686 AACAGACGCTGAGCAGAGATGTTCCGGCGCTCTTGCCATGCGCAAGCTGCGCAGAG 745
QY 201 MetSerProAspLeuHisArgGlyValSerValGlnTyrcysAspValLeuHisLysLys 220
DB 746 ATGAGCCAGACCTGCAACCGCAAGTCTCGGTGCAAGTGCAGCGTGCACCAAGAG 805
QY 221 AlaLeuArgAsnLysLysLeuLeuArgAlaGlySerGlyGlyGlyGlyAspProGly 240
DB 806 CGCTCGGAGAACAAAGAGCTGCTCGGCGCGGAGCGGCGGCGGCGGCGGCGGCGG 865
QY 241 AspAlaPheGlyIleValAlaProPheAlaArgArgProSerValHisSerAspLeuMet 260
DB 866 GACGCTTGGCATGTCGTGGACCTTCGCGCGCGGCGGCGGCGGCGGCGGCGGCGG 925
QY 261 TyrlleArgGlyLysAlaSerAlaGlySerGlnAlaLysAspLysGluArgCysValIle 280
DB 926 TACATCCCGGAGAGGCGGAGCGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCG 985
QY 281 Ser 281
DB 986 AGC 988
RESULT 3
ID AA236893 standard; cDNA; 1801 BP.
XX AC AA236893;
XX AC AA236893;
DT 13-MAR-2000 (first entry)
XX DE cDNA encoding an activator of G protein signalling (AGS) protein.
XX KW Activator of G protein signalling; AGS; ras-related G protein;
XX KW GTP hydrolysis; G protein activity; pheromone response pathway;
KW G protein-coupled signal transduction; G-gamma selectivity;
cellular signal transduction; ss.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
FT 5'UTR 7..153
FT CDS /*tag= a
FT 154..999
FT /*tag= b
FT /product= "activator of G protein signalling (AGS)
FT 3'UTR 1000..1801
FT /*tag= c
XX PN W09958670-A1.
XX PD 18-NOV-1999.
XX PF 07-MAY-1999; 99WO-US10151.
XX PR 08-MAY-1998; 98US-0084842.
XX PR 07-OCT-1998; 98US-0103355.
XX PA (CADU-) CADUS PHARM CORP.
XX PI Cismowski M, Duzic E;
XX DR MPI: 2000-072337/06.
XX DR P-PSDB; AAY53921.
XX PT A new activator of G protein signalling used to treat disorders
XX PT characterized by an aberrant AGS protein activity -
PS Claim 3; Page 133-135; 162pp; English.

CC The present sequence encodes an activator of G protein signalling (AGS)
 CC protein. The cDNA sequence was isolated from a human liver cDNA
 CC library. The AGS protein exhibits homology to ras-related G proteins,
 CC and contains alterations in conserved amino acids consistent with a
 CC deficiency in GTP hydrolysis activity. AGS stimulates G protein
 CC activity, G protein-coupled signal transduction and the pheromone
 CC response pathway in a receptor-independent manner. The AGS protein
 CC also shows G-gamma selectivity, as measured by growth assays in
 CC yeast expressing various mammalian G-gamma constructs, and
 CC tissue-specific expression, as measured by Northern blot analysis.
 CC The AGS protein can be used to screen for compounds that modulate
 CC cellular signal transduction. The protein is used to treat disorders
 CC characterized by an aberrant AGS protein activity or AGS nucleic acid
 CC expression.
 XX

SQ Sequence 1801 BP; 437 A; 561 C; 500 G; 303 T; 0 other;

Alignment Scores:
 Pred. No.: 2.02e-157 Length: 1801
 Score: 1449.00 Matches: 281
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 99.93% Indels: 0
 DB: 21 Gaps: 0

US-09-709-103-1F1 (1-282) x AAZ236893 (1-1801)

QY 1 MetLysLeuAlaAlaMetIleLysLysMetCysProSerAspSerGluLeuSerIlePro 20
 DB 154 ATGAACCTGGCGCGATGATCAAGAAGATGTCGCCGAGCGACTCGAGTATCCCG 213
 QY 21 AlaLysAsnCysTyrArgMetValIleLeuGlySerSerLysValGlyLysThrAlaIle 40
 DB 214 GCCAAGAAGTCTATCGCATGCTATCTCGCTCGTCCAAAGGTGGCAAGCGGCATC 273
 QY 41 ValSerArgPheLeuThrGlyArgPheGluAspAlaTyrThrProThrIleGluAspPhe 60
 DB 274 GTGTCGCGCTTCTCCACCGCGCGTTCGAGGACGCTACAGCGCTACCATCGAGACTTC 333
 QY 61 HisArgLysPheTyrSerIleArgGlyGluValTyrGlnLeuAspIleLeuAspThrSer 80
 DB 334 CACCGCAAGTTCTACTCCATCCGCGGAGGTCTACCAGCTCGAGCATCTCCGACACGTCC 393
 QY 81 GlyAsnHisProPheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIle 100
 DB 394 GGCACACACCGTTCCTCCGCGCTCGCGGCTCTCCATCTCAGAGAGAGTTTTCATC 453
 QY 101 LeuValPheSerLeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGlnGln 120
 DB 454 CTGGTGTTCAGTCTGGACAACCGCGACTCTCTTCGAGGAGGTGCGAGCGCTCAGCAGCAG 513
 QY 121 IleLeuAspThrLysSerCysLeuLysAsnLysThrLysGluAsnValAspValProLeu 140
 DB 514 ATCTCTGACACCAAGTCTTGGCTCAAGAACAAACCAAGGAGAACGTGGACGTGCCCTG 573
 QY 141 ValIleCysGlyAsnLysGlyAspArgPheTyrArgGluValAspGlnArgGluIle 160
 DB 574 GTCATCTCGGCAACAAGGGTACCGGACTTCTACCGGAGGTGGACCGCGAGATC 633
 QY 161 GluGlnLeuValGlyAspProGlnArgCysAlaTyrPheGluIleSerAlaLysLys 180
 DB 634 GAGCAGCTGTGTGGGCGAGACCCCGCGCTGCGCTTCTCGAGATCTCGGCCAAGAAG 693
 QY 181 AsnSerSerLeuAspGlnMetPheArgAlaLeuPheAlaMetAlaLysLeuProSerGlu 200
 DB 694 AACAGACGCTGGACCAAGATCTTCGCGCGCTCTTCGCCATGGCCAAAGCTGCCAGCGAG 753
 QY 201 MetSerProAspLeuHisArgLysValSerValGlnTyrCysAspValLeuHisLysLys 220
 DB 754 ATGAGCCAGACCTTGCACCGCAAGGTCTCGGTGCAGTACTGCCAGCGTCTGCACAGAAG 813
 QY 221 AlaLeuArgAsnLysLysLeuLeuArgAlaGlySerGlyGlyGlyGlyAspProGly 240

DB 814 GCGCTCGGGAACAAGAAGCTGCTCGGGCGCGCAGCGCGCGCGCGACCCGGGC 873
 QY 241 AspAlaPheGlyIleValIalaProPheAlaArgArgProSerValHisSerAspLeuMet 260
 DB 874 GAGCGCTTTGGCATCTGTGCACCCCTTCGGCGCGCGCGCGCGGTACACAGCGACCTCATG 933
 QY 261 TyrIleArgGluLysAlaSerAlaGlySerGlnAlaLysAspLysGluArgCysValIle 280
 DB 934 TACATCCGCGAGAAGCCAGCCGCGCGCGCGCGCGCAAGGACCAAGAGCGCTGCGCTCATC 993
 QY 281 Ser 281
 DB 994 AGC 996
 RESULT 4
 AAZ23022
 ID AAZ23022 standard; cDNA; 1841 BP.
 XX
 AC AAZ23022;
 XX
 DT 17-JAN-2000 (first entry)
 XX
 DE Human kd312 polypeptide encoding cDNA.
 XX
 KW kd312 polypeptide; cancer; immunodeficiency disorder; AIDS; stroke;
 KW heart attack; head trauma; neurodegenerative disease; human;
 KW Parkinson's disease; Alzheimer's disease; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO950288-A2.
 XX
 PD 07-OCT-1999.
 XX
 PF 30-MAR-1999; 99WO-US06993.
 XX
 PR 31-MAR-1998; 98US-0053374.
 XX
 PA (AMGE-) AMGEN INC.
 XX
 PI Yen K;
 XX
 WPI: 1999-601322/51.
 P-PSDB; AAY42693.
 XX
 PT kd312 polypeptides useful for treating diseases and disorders
 associated with alterations in cell proliferation and cell death
 XX
 PS Claim 1; Fig 8; 85pp; English.
 XX
 CC The invention provides nucleic acid molecules encoding human and rat
 kd312 polypeptides. The kd312 polypeptides can be expressed by standard
 recombinant methodology. The kd312 sequences, and the antibodies against
 the proteins may be used to treat or diagnose the presence or progression
 of various diseases, e.g. cancer, immunodeficiency disorders (e.g. AIDS),
 CC stroke, heart attack, head trauma, and neurodegenerative diseases (e.g.
 Parkinson's disease and Alzheimer's disease). The present sequence
 CC represents the human kd312 cDNA sequence.
 XX
 SQ Sequence 1841 BP; 398 A; 605 C; 530 G; 308 T; 0 other;
 Alignment Scores:
 Pred. No.: 2.09e-157 Length: 1841
 Score: 1449.00 Matches: 281
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 99.93% Indels: 0
 DB: 20 Gaps: 0
 US-09-709-103-1F1 (1-282) x AAZ23022 (1-1841)
 QY 1 MetLysLeuAlaAlaMetIleLysLysMetCysProSerAspSerGluLeuSerIlePro 20

Db 255 ATGAACTGCGCCGATGATCAAGAAATGTCGCCGAGCTCGAGCTGATATCCCG 314
 QY 21 AAllysasnCyseTyArsmetValilleuGlyseSerSeTyValGlyeThzAlle 40
 Db 315 GCCAAGAACTGCTATGCGATGTCATCTCCGCTCGTCCAAAGTGGCAAGCGCCATC 374
 QY 41 ValSerArgPheLeuThrGlyArgPheGluAspAlaTyThrProThrIleGluasphe 60
 Db 375 GTGTGCGCTTCTCTCCACCGCGCTTCGAGGACGCTTACACCTTACATCGAGGACTTC 434
 QY 61 HlsArgLysPheThySerIleArgGlyGluValTyrgLInleuAspIleleuaspThrSer 80
 Db 435 CACCGCAAGTCTACTCATTCGCGCGAGGTCTCAAGCTCGACATCTCCACACGCTCC 494
 QY 81 GlysAsnHisProPheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIle 100
 Db 495 GGCAACCAACCGCTTCCCGCCCATGGCGGCCCTCTCCATCTCACAGAGAGCTTTTCATC 554
 QY 101 LeuValPheSerLeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGlnGln 120
 Db 555 CTGTGTTCTAGTCTGACAAACCGCAGCTCTTCGAGAGGTGACAGCTCAGCGACGAG 614
 QY 121 IleLeuAspThyLeuSerCysLeuLysAsnLysThrLysGluAsnValAspValProLeu 140
 Db 615 ATCTTGACACCAAGCTTGTCTCAAGAAACAAACCAAGAGACGTGACGCTGCCCTG 674
 QY 141 ValIleCysGlyAsnLysGlyAspArgPheTyArgGlyValAspGlnArgGlnIle 160
 Db 675 GTCATCTGGCGCAACAGGTGACCCGCACTTCTACCGCGAGGTGACCAAGCGCAGATC 734
 QY 161 GluGlnLeuValGlyAspAspProGlnArgCysAlaTyThrPheGluIleSerAlaLysLys 180
 Db 735 GAGCAGCTGTGGCGACGACCCCGCAGCTGCGCTTACTGAGATCTCGGCCAAGAG 794
 QY 181 AsnSerSerLeuAspGlnMetPheArgAlaLeuPheAlaMetAlaLysLeuProSerGln 200
 Db 795 AACAGAGCTGTGACCAAGTGTTCGCGCGCTTCTCGCATGCGCAAGCTGCCCGCAGAG 854
 QY 201 MetSerProAspLeuHisArgLysValSerValGlnTyThrCysAspValLeuHisLysLys 220
 Db 855 ATGAGCCCGACAGCTGCAACCGCAAGTCTGTGCGAGTACTGCGAGCTGTGCGACAAAG 914
 QY 221 AlaLeuArgAsnLysLysLeuLeuArgAlaGlySerGlyGlyGlyGlyAspProGly 240
 Db 915 GCGCTGCGGAACAAAGAGCTGTGCGCGCGCGGCGAGCGCGCGCGCGACCGCGGCG 974
 QY 241 AspAlaPheGlyIleValAlaProPheAlaArgArgProSerValHisSerAspLeuMet 260
 Db 975 GACGCTTGTGATGCTGCGACCTTCCGCGCGCCCGCCACAGCTACACAGCACTCATG 1034
 QY 261 TyrlleArgGluLysAlaSerAlaGlySerGlnAlaLysAspLysGluArgCysValIle 280
 Db 1035 TACATTCGCGAAGAGCCAGCGCGCGCGACCGCCAGCCAGCAAGAGAGCGCTCATC 1094
 QY 281 Ser 281
 Db 1095 AGC 1097
 RESULT 5
 AAA49177
 ID AAA49177 standard; cDNA, 1776 BP.
 XX
 AC AAA49177;
 XX
 DT 03-NOV-2000 (first entry)
 XX
 DE cDNA encoding human GTPase associated protein-7.
 XX
 KW Guanine nucleotide binding protein; GMP-binding protein; G-protein;
 KW GTPase; GTPase associated protein; GTPAP; cell proliferation;
 KW autoimmune; inflammatory; immune system disorder; cancer; AIDS;
 KW acquired immune deficiency syndrome; asthma; atherosclerosis;
 KW arthritis; systemic lupus erythematosus; psoriasis; human; ss.

XX OS Homo sapiens.
 XX Location/Qualifiers
 FH Key 180..1025
 FT CDS
 FT /tag= a
 FT /product= GTPAP7
 XX MO200031263-A2.
 XX
 PD 02-JUN-2000.
 XX
 PF 23-NOV-1999; 99MO-US28013.
 XX
 PR 23-NOV-1998; 98US-0109592.
 PR 04-FEB-1999; 99US-0118610.
 PR 06-APR-1999; 99US-0127990.
 XX
 PA (INCY-) INCYTE PHARM INC.
 XX
 PI Hillman JL, Tang YT, Bandman O, Lal P, Yue H, Lu DAM, Baughn MR;
 PI Yang J, Azimzai Y;
 XX
 DR WPI; 2000-400073/34.
 DR P-PSDB; AAY99655.
 XX
 PT Human GTPase associated proteins, polynucleotides, and antibodies.
 PT useful for diagnosing, preventing and treating various diseases such as
 PT atherosclerosis, cancer, acquired immune deficiency syndrome (AIDS),
 PT asthma, and autoimmune diseases -
 XX
 PS Claim 9; Page 125-126; 144p; English.
 XX
 CC Human cDNA libraries from various tissues were screened for GTPase
 CC associated proteins (GTPAP). The present sequence is cDNA encoding
 CC human GTPAP-7. This sequence was derived from a cDNA library of the
 CC brain tumour tissue from the parietal lobe of a female.
 CC This protein is expressed in reproductive, nervous and
 CC gastrointestinal tissue. The GTPAP proteins may be used to define
 CC agonists and antagonists of GTPAP activity and to generate antibodies
 CC to GTPAP. This means the GTPAP proteins may be useful for treatment or
 CC prevention of diseases associated with GTPAP such as cell proliferation
 CC disorders, autoimmune disorders, inflammatory disorders, immune system
 CC disorders, cancer, AIDS, asthma, atherosclerosis, arthritis, systemic
 CC lupus erythematosus and psoriasis.
 XX
 SQ Sequence 1776 BP; 430 A; 565 C; 490 G; 291 T; 0 other;
 Alignment Scores:
 Pred. No.: 1.67e-156 Length: 1776
 Score: 1441.00 Matches: 280
 Percent Similarity: 99.64% Conservative: 0
 Best Local Similarity: 99.64% Mismatches: 1
 Query Match: 99.38% Indels: 0
 DB: 21 Gaps: 0
 US-09-709-103-1f1 (1-282) x AAA49177 (1-1776)
 QY 1 MetLysLeuAlaAlaMetIleLysLysMetCysProSerAspSerGluLeuSerIlePro 20
 Db 180 ATGAACTGCGCCGATGATCAAGAAATGTCGCCGAGCTCGAGCTGATATCCCG 239
 QY 21 AAllysasnCyseTyArsmetValilleuGlyseSerSeTyValGlyeThzAlle 40
 Db 240 GCCAAGAACTGCTATGCGATGTCATCTCCGCTCGTCCAAAGTGGCAAGCGGCCATC 299
 QY 41 ValSerArgPheLeuThrGlyArgPheGluAspAlaTyThrProThrIleGluasphe 60
 Db 300 GTGTGCGCTTCTCTCCACCGCGCTTCGAGGACGCTTACATCGAGGAGCTTC 359
 QY 61 HlsArgLysPheThySerIleArgGlyGluValTyrgLInleuAspIleleuaspThrSer 80
 Db 360 CACCGCAAGTCTACTCATTCGCGCGAGGTCTACAGCTCGACATCTCCGACACGCTCC 419

Qy 81 GlyAsnHisProPheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIle 100
 Db 420 GGCAACACCGCTTCCCGCCATCGGTGCTCTCCATCTCCACAGAGAGCGTTTTCATC 479
 Qy 101 LeuValPheSerLeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGlnGln 120
 Db 480 CTGGTGTTCAGTCTGGACAAACCGGACTCTCTCGAGAGAGGTGACGGCTCAGGCGAGCAG 539
 Qy 121 IleLeuAspThrLysSerCysLeuLysAsnLysThrLysGluAsnValAspValProLeu 140
 Db 540 ATCTCTGACACCAAGTCTTGCTCTCAAGACAAACACCAAGAGAGCGTGAGCGTCCCTG 599
 Qy 141 ValIleCysGlyAsnLysGlyAspArgAspPheTyrArgGluValAspGlnArgGluIle 160
 Db 600 GTCATCTGCGGCAACAAGGCTGACCGGACTTCTTACCGCGAGGTGGACGAGCGGAGATC 659
 Qy 161 GluGlnLeuValGlyAspAspProGlnArgCysAlaTyrPheGluIleSerAlaLysLys 180
 Db 660 GAGCAGCTGGTGGCGGACGACCCGACGCTGCGCCTTACTTCGAGATCTCGGCCAAGAAG 719
 Qy 181 AsnSerSerLeuAspGlnMetPheArgAlaLeuPheAlaMetAlaLysLeuProSerGlu 200
 Db 720 AACAGCAGCTGGACAGATGTTCCGCGCGTCTTCGCCATGGCCAGACTGCCAGCGAG 779
 Qy 201 MetSerProAspLeuHisArgLysValSerValGlnTyrCysAspValLeuHisLysLys 220
 Db 780 ATGAGCCAGACCTGCACCGCAAGGTCTCGGTGCGAGTACTGCGACGTGCTGCACAAGAG 839
 Qy 221 AlaLeuArgAsnLysLysLeuLeuArgAlaGlySerGlyGlyGlyGlyAspProGly 240
 Db 840 GCGTTCGGGAACAAGCTCTCGCGCGCGGACGCGCGCGCGCGCGCGCGCGCGCGCGCG 899
 Qy 241 AspAlaPheGlyIleValAlaProPheAlaArgArgProSerValHisSerAspLeuMet 260
 Db 900 GACGCTTTGGCATCGTGGCACCTCTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 959
 Qy 261 TyrIleArgGluLysAlaSerAlaGlySerGlnAlaLysAspLysGluArgCysValIle 280
 Db 960 TACATCCGCGAAGAGCG 1019
 Qy 281 Ser 281
 Db 1020 AGC 1022
 RESULT 6
 AA223024
 ID AA223024 standard; cDNA; 1689 BP.
 XX
 AC AA223024;
 XX
 DT 17-JAN-2000 (first entry)
 XX
 DE Rat kd312 polypeptide encoding cDNA.
 XX
 KW kd312 polypeptide; cancer; immunodeficiency disorder; AIDS; stroke;
 KW heart attack; head trauma; neurodegenerative disease; rat;
 KW Parkinson's disease; Alzheimer's disease; ss.
 XX
 OS Rattus sp.
 XX
 PN W09950288-A2.
 XX
 PD 07-OCT-1999.
 XX
 PF 30-MAR-1999; 99WO-US06993.
 XX
 PR 31-MAR-1998; 98US-0053374.
 XX
 PA (AMGE-) AMGEN INC.
 XX
 PI Yen K;
 XX

DR WPI: 1999-601322/51.
 DR P-PSDB; AAY42694.
 XX
 PT kd312 polypeptides useful for treating diseases and disorders
 PT associated with alterations in cell proliferation and cell death -
 XX
 PS Claim 2; Fig 9; 85pp; English.
 XX
 CC The invention provides nucleic acid molecules encoding human and rat
 CC kd312 polypeptides. The kd312 polypeptides can be expressed by standard
 CC recombinant methodology. The kd312 sequences, and the antibodies against
 CC the proteins may be used to treat or diagnose the presence or progression
 CC of various diseases, e.g. cancer, immunodeficiency disorders (e.g. AIDS),
 CC stroke, heart attack, head trauma, and neurodegenerative diseases (e.g.
 CC Parkinson's disease and Alzheimer's disease). The present sequence
 CC represents the rat kd312 cDNA sequence.
 XX
 SQ Sequence 1689 BP; 417 A; 462 C; 434 G; 376 T; 0 other;

Alignment Scores:
 Pred. No.: 9.87e-152 Length: 1689
 Score: 1399.50 Matches: 273
 Percent Similarity: 98.22% Conservative: 3
 Best Local Similarity: 97.15% Mismatches: 4
 Query Match: 96.52% Indels: 1
 Gaps: 20
 DB:

US-09-709-103-1F1 (1-282) x AA223024 (1-1689)
 Qy 1 MetLysLeuAlaAlaMetIleLysLysMetCysProSerAspSerGluLeuSerIlePro 20
 Db 132 ATGAACCTGGCGCGCATGATCAAGAAGATGTGCCAAGACTCTGAACCTGAGTATCCCG 191
 Qy 21 AlaLysAsnCysTyrArgMetValIleLeuGlySerSerLysValGlyLysThrAlaIle 40
 Db 192 GCCAAGACTGCTACAGATGGTTCATCTCGGCTCATCAAGTGGGCAAGCGCCATC 251
 Qy 41 ValSerArgPheLeuThrGlyArgPheGluAspAlaTyrThrProThrIleGluAspPhe 60
 Db 252 GTGTGCGGCTTCTCACGGCGGCTTCGAGGAGCGCTTACACCCCTACCATTTGAAGACTTC 311
 Qy 61 HisArgLysPheTyrSerIleArgGlyGluValTyrGlnLeuAspIleLeuAspThrSer 80
 Db 312 CACCGAAGATTTTACTCGATCCGCGGCAAGTCTACCAAGTTGGACATACCTGGACACATCT 371
 Qy 81 GlyAsnHisProPheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIle 100
 Db 372 GGCAATCATCTCGTTTCCCGCCATCGCGGCTCTCTATCTCTCACAGGAGCGTTTTCATT 431
 Qy 101 LeuValPheSerLeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGlnGln 120
 Db 432 CTGGTGTTCAGCTTAGCAACCGCGACTCTTCGAGGAGGTGCAAGGCTCAACACGAG 491
 Qy 121 IleLeuAspThrLysSerCysLeuLysAsnLysThrLysGluAsnValAspValProLeu 140
 Db 492 ATCTTAGACCAACAGTCTCTCTCAAGAACAAACCAAGAGAAATGTGGAGCGTCCGCTG 551
 Qy 141 ValIleCysGlyAsnLysGlyAspArgAspPheTyrArgGluValAspGlnArgGluIle 160
 Db 552 GTCATTTGGGTAAACAAGGGGACCGGACTTCTACCCGGAAGTGGAGCAGCGGAGATT 611
 Qy 161 GluGlnLeuValGlyAspAspProGlnArgCysAlaTyrPheGluIleSerAlaLysLys 180
 Db 612 GAGCAGCTGGTGGCGGATGACCCCTCAGCGTTTGTGCTACTTCGAGATCTCGGCCAAGAAG 671
 Qy 181 AsnSerSerLeuAspGlnMetPheArgAlaLeuPheAlaMetAlaLysLeuProSerGlu 200
 Db 672 AATAGCAGCTTGACACAGATGTTCCGTGCGCTCTTTGCGATGGCCCAAGCTCGCTAGCGAG 731
 Qy 201 MetSerProAspLeuHisArgLysValSerValGlnTyrCysAspValLeuHisLysLys 220
 Db 732 ATGAGCCCTGACTTGCACCGCAAGGTGCTGTGCAAGTACTGTGACGTCTGTCACAAAAAG 791

QY 221 AlaLeuArgAsnLysLysLeuLeuAlaGlySerGlyGlyGlyGlyAspProGly 240
 Db 792 GCTCTGGAGCAAGAGAGCTTCTGCGCGGCGAGC---GAGAGTGGGGCGACCAAGCA 848
 QY 241 AspAlaPheGlyIleValAlaProPheAlaArgArgProSerValHisSerAspLeuMet 260
 Db 849 GATGCTTTGGCATTTGGCGCCCTTGTCTGCGAGCTAGCGCTGATAGCGACCTCATG 908
 QY 261 TyrIleArgGluLysAlaSerAlaGlySerGlnAlaLysAspLysGluArgCysValIle 280
 Db 909 TACATTCGTAGAAAACCACTGTCAGAGAGCCAGGCTTAGGACAGAGAGCGCTGTGATC 968
 QY 281 Ser 281
 Db 969 AGT 971

RESULT 7
AA223023

ID AA223023 standard; DNA; 3986 BP.

XX AA223023;

DT 17-JAN-2000 (first entry)

DE Human kd312 genomic DNA sequence.

XX kd312 polypeptide; cancer; immunodeficiency disorder; AIDS; stroke;
 KW heart attack; head trauma; neurodegenerative disease; human;
 KW Parkinson's disease; Alzheimer's disease; ss.

OS Homo sapiens.

XX WO950288-A2.

XX PD 07-OCT-1999.

XX PF 30-MAR-1999; 99WO-US06993.

XX PR 31-MAR-1998; 98US-0053374.

XX PA (AMGE-) AMGEN INC.

XX PI Yen K;

XX DR WPI; 1999-601322/51.

XX DR P-PsDB; AAY42693.

XX PT kd312 polypeptides useful for treating diseases and disorders
 associated with alterations in cell proliferation and cell death

XX PS Claim 1; Fig 5; 85pp; English.

XX The invention provides nucleic acid molecules encoding human and rat
 CC kd312 polypeptides. The kd312 polypeptides can be expressed by standard
 CC recombinant methodology. The kd312 sequences, and the antibodies against
 CC the proteins may be used to treat or diagnose the presence or progression
 CC of various diseases, e.g. cancer, immunodeficiency disorders (e.g. AIDS),
 CC stroke, heart attack, head trauma, and neurodegenerative diseases (e.g.
 CC Parkinson's disease and Alzheimer's disease). The present sequence
 CC represents the human kd312 genomic DNA sequence.

XX SQ Sequence 3986 BP; 856 A; 1226 C; 1172 G; 732 T; 0 other;

Alignment Scores:

Pred. No.:	9.366-150	Length:	3986
Score:	1387.00	Matches:	281
Percent Similarity:	79.83%	Conservative:	0
Best Local Similarity:	79.83%	Mismatches:	0
Query Match:	95.66%	Indels:	71
DB:	20	Gaps:	1

US-09-709-103-1f1 (1-282) X AA223023 (1-3986)

QY 1 MetLysLeuAlaAlaMetIleLysLysMetCysProSerAspSerGluLeuSerIlePro 20
 Db 774 ATGAAACCTGGCGCGGATGATCAAGAGATGCGCCGAGCACTGGAGCTGATGATCCG 833
 QY 21 AlaLysAsnCysTyrArgMetValIleLeuGlySerSerLysValGlyLysThrAlaIle 40
 Db 834 GCCAAGAACTGCTATCCATGATGATGATCTCGGCTGCTCCAGGAGGCAAGACGGCCATC 893
 QY 41 ValSerArgPheLeuThrGlyArgPheGluAspAlaTyrThrProThrIleGluAspPhe 60
 Db 894 GTGTGGCGCTTCTCACCAGCGCGCTTGAGAGCGCTTACACGCTTACATCAGAGACTTC 953
 QY 61 HisArgLysPheTyrSerIleArgGlyIleValTyrGlnLeuAspIleLeuAspThrSer 80
 Db 954 CACCGCAAGTTCTACTCATTCGCGGCGAGGTCTTACAGCTTCAGATCTTCGACAGCTTC 1013
 QY 81 GlyAsnHisProPheProAlaMetArgArgLeuSerIleLeuThr----- 95
 Db 1014 GGCAACCAACCCGCTTCCCGCATGCGCGCTCTCCATCTCACAGGTGAGCCGGGGGCC 1073
 QY 95 ----- 95
 Db 1074 GGGCAGGTGCGGAGGAGGAGGCGGGGAACTTCGCGCAGGCGCCCGAGCGCGGT 1133
 QY 95 ----- 95
 Db 1134 CCGGCTGCTGCGCGCGAGTAGTACGCTTGCGCTTAGAGAGCTAGCGGCCCGCGCG 1193
 QY 95 ----- 95
 Db 1194 GCTCAAAAGTCAGCCGACTTGTCCCTGGCGGCGCACTTCACCTTCTCTTTCGCT 1253
 QY 96 -----GlyAspValPheIleLeuValPheSerLeuAspAsnArgAspSe 110
 Db 1254 CTCTGTCCCCCTTCTAGAGAACGTTTTCATCTGCTGCTTCACTGAGCAACCGCATCTC 1313
 QY 110 rPheGluGluValGlnArgLeuArgGlnGlnIleLeuAspThrLysSerCysLeuLysAs 130
 Db 1314 CTTGAGAGAGGTGAGCGGCTCAGGACGACGATCTCGACACCAAGCTTTCCTCAAGAA 1373
 QY 130 rLysThrLysGluAsnValAspValProLeuValIleCysGlyAsnLysGlyAspArgAs 150
 Db 1374 CAAGAACCAAGAGAACGTCGACGCTGCCCTGTCATCTGGCGCAACAGGTCACCGCA 1433
 QY 150 rPheTyrArgGluValAspGlnArgGluIleGluGlnLeuValGlyAspAspProGlnr 170
 Db 1434 CTTCTACCGCGAGGTGACCAAGCGCAGATCGAGCTGTGTGGCGAGCAACCCACAGCG 1493
 QY 170 rCysAlaTyrPheGluIleSerAlaLysLysAsnSerSerLeuAspGlnMetPheArgAl 190
 Db 1494 CTGGCGCTTACTTTCGAGATCTCGGCGCAAGAAAGACAGACGCTTGACCAAGTGTTCGCGCG 1553
 QY 190 rLeuPheAlaMetAlaLysLeuProSerGluMetSerProAspLeuHisArgLysValSe 210
 Db 1554 GCTCTTCGCGATGCGCAAGCTGCGCCAGCAGATGAGCAGACCTTGACCGCAAGGTCCTC 1613
 QY 210 rValGlnTyrCysAspValLeuHisLysLysAlaLeuArgAsnLysLysLeuLeuArgAl 230
 Db 1614 GGTGCACTACTGAGACGTGTGCAAGAGAGCGCTGGCAACAAAGACTGTGGGGGC 1673
 QY 230 rGlySerGlyGlyGlyGlyGlyAspProGlyAspAlaPheGlyIleValAlaProPheAl 250
 Db 1674 CGGAGCGCGCGCGCGCGCGCGCGACCCCGGCGACGCTTGGCATGTGTGACCTTCCG 1733
 QY 250 rArgArgProSerValHisSerAspLeuMetTyrIleArgGluLysAlaSerAlaGlySe 270
 Db 1734 GCGCGCGCGCGCGAGGTGACAGCAGCACTTCATGTACATCCCGAGAGAGCGCGCGCGAG 1793
 QY 270 rGlnAlaLysAspLysGluArgCysValIleSer 281
 Db 1794 CAGGCCAAGAGCAAGAGAGCGCTGCTCATTCAGC 1827

RESULT 8

```

AAZ23025
XX AAZ23025 standard; DNA; 3079 BP.
XX AC AAZ23025;
XX DT 17-JAN-2000 (first entry)
XX XX Rat kd312 genomic DNA sequence.
XX KW kd312 polypeptide; cancer; immunodeficiency disorder; AIDS; stroke;
XX KW heart attack; head trauma; neurodegenerative disease; rat;
XX KW Parkinson's disease; Alzheimer's disease; ss.
XX OS Rattus sp.
XX XX WO950288-A2.
XX PN 07-OCT-1999.
XX PF 30-MAR-1999; 99WO-US06993.
XX PR 31-MAR-1998; 98US-0053374.
XX XX (AMGE-) AMGEN INC.
XX PA Yen K;
XX PI WPI; 1999-601322/51.
XX DR P-PSDB; AAY42694.
XX XX kd312 polypeptides useful for treating diseases and disorders
XX PT associated with alterations in cell proliferation and cell death -
XX PS Claim 2; Fig 7; 85pp; English.
XX CC The invention provides nucleic acid molecules encoding human and rat
XX CC kd312 polypeptides. The kd312 polypeptides can be expressed by standard
XX CC recombinant methodology. The kd312 sequences, and the antibodies against
XX CC the proteins may be used to treat or diagnose the presence or progression
XX CC of various diseases, e.g. cancer, immunodeficiency disorders (e.g. AIDS),
XX CC stroke, heart attack, head trauma, and neurodegenerative diseases (e.g.
XX CC Parkinson's disease and Alzheimer's disease). The present sequence
XX CC represents the rat kd312 genomic DNA sequence.
XX SQ Sequence 3079 BP; 708 A; 843 C; 829 G; 699 T; 0 other;

Alignment Scores:
Pred. No.: 7,99e-145 Length: 3079
Score: 1343.00 Matches: 272
Percent Similarity: 84.10% Conservative: 3
Best Local Similarity: 83.18% Mismatches: 5
Query Match: 92.62% Indels: 48
DB: 20 Gaps: 2

us-09-709-103-1f1 (1-282) x AAZ23025 (1-3079)
Qy 1 MetLysLeuAlaAlaMetIleLysLysMetCysProSerAspSerGluLeuSerIlePro 20
Db 303 ATGAACTGGCGGATGATCAAGAAGATGTGCCAAGCGACTCTGAAGTATCCCG 362
Qy 21 AlaLysAsnCysTyrArgMetValIleLeuGlySerSerLysValGlyLysThrAlaIle 40
Db 363 GCCAAGAATCTGTACAGATGGTATCTCGGCTCATCCAAAGTGGGCAAGCGGCATC 422
Qy 41 ValSerArgPheLeuThrGlyArgPheGluAspAlaTyrThrProThrIleGluAspPhe 60
Db 423 GTGTGCGCTTCTCTCACGGCGGCTTCAGGACGCTTACACCCCTACCATTTGAAGACTTC 482
Qy 61 HisArgLysPheTyrSerIleArgGlyGluValTyrGlnLeuAspIleLeuAspThrSer 80
Db 483 CACCCGAAAGTTTACTCGATCCGCGGGAAGTCTACCAAGTTGGACATACCTGGACATCT 542
Qy 81 GlyAsnHisProPheProAlaMetArgLeuSerIle----- 93

AAZ23025
Db 543 GGCAATCATCGTTTCCCGCATCGCGGGCTCTCTATCTCACAGTGAGTGGGGACC 602
Qy 93 ----- 93
Db 603 GACAGGACCGTGGGAGGAATCTCGGGAGCGGATGGGGCGGTGTGTGTGCTTTGGG 662
Qy 94 -----Leu 94
Db 663 GCTGTGCTGTCTGCTGCTCCGTGCTGGCAGTCCCTCACCTTTCCACTGCTTCCCTTG 722
Qy 95 ThrGlyAspValPheIleLeuValPheSerLeuAspAsnArgAspSerPheGluGluVal 114
Db 723 TA-GGAGACGTTTTCATCTCTGTTTCTAGCAACCGGACTCTCTCGAGAGGTG 781
Qy 115 GlnArgLeuArgGlnGlnIleLeuAspThrLysSerCysLeuLysAsnLysThrLysGlu 134
Db 782 CAAAGGCTCAACACAGCAGATCTTAGACACCAAGTCTCTCAAGAACAAACCAAGAG 841
Qy 135 AsnValAspValProLeuValIleCysGlyAsnLysGlyAspArgAspPheTyrArgGlu 154
Db 842 AATGTGGACGTGCCGCTGGTCAATTGGCGGTAAACAAAGGGGACCGGACTTCTACCGGAA 901
Qy 155 ValAspGlnArgGluIleGluGlnLeuValGlyAspAspProGlnArgCysAlaTyrPhe 174
Db 902 GTGGAGCAGCGGAGATTGACAGCTGGTGGCGCATGCCCTCAGCGTTGTCCTACTTC 961
Qy 175 GluIleSerAlaLysLysAsnSerSerLeuAspGlnMetPheArgAlaLeuPheAlaMet 194
Db 962 GAGATCTCGGCCAAGAGATAGCAGCTGGTGGCGATGTTCCGTCGCTCTTTGCCATG 1021
Qy 195 AlaLysLeuProSerGluMetSerProAspLeuHisArgLysValSerValGlnTyrCys 214
Db 1022 GCCAAGCTGCCTAGCAGATGAGCCCTGACTTGCCCAAGGTGTCTGTGCTGCTACTGT 1081
Qy 215 AspValLeuHisLysLysAlaLeuArgAsnLysLysLeuLeuArgAlaGlySerGlyGly 234
Db 1082 GACGTGCTGCACAAAGAGCTCTGAGGAACAAAGAGCTTCTGCGTGGCGCAGC--GGA 1138
Qy 235 GlyGlyGlyAspProGlyAspAlaPheGlyIleValAlaProPheAlaArgArgProSer 254
Db 1139 CGTGGGGCGCACCGAGATGCCCTTTGGCATCTTGGCGCCCTTTGCTCGCAGACCTAGC 1198
Qy 255 ValHisSerAspLeuMetTyrIleArgGluLysAlaSerAlaGlySerGlnAlaLysAsp 274
Db 1199 GTGCATAGCGACCTCATGTATCTGTGAGAAACCAAGTGTCTGTCAGCAGCTAAGGAC 1258
Qy 275 LysGluArgCysValIleSer 281
Db 1259 AAGGAGCGCTGTGTATCAGT 1279
RESULT 9
AAZ36913
ID AAZ36913 standard; DNA; 837 BP.
XX AC AAZ36913;
XX XX AAZ36913;
XX DT 13-MAR-2000 (first entry)
XX DE DNA encoding a homologue of activator of G protein signalling AGS1.
XX KW Activator of G protein signalling; AGS; ras-related G protein;
XX KW GTP hydrolysis; G protein activity; pheromone response pathway;
XX KW G protein-coupled signal transduction; G-gamma selectivity;
XX KW cellular signal transduction; AGS1 homologue; ss.
XX OS Homo sapiens.
XX XX Location/Qualifiers
XX FH Key 1..837
XX FT CDS /*tag= a
XX FT /product= "AGS1 homologue"
XX XX

```

PN MO9958670-A1.
 XX 18-NOV-1999.
 PD
 XX 07-MAY-1999; 99WO-US10151.
 PF
 XX 08-MAY-1998; 98US-0084842.
 PR 07-OCT-1998; 98US-0103355.
 XX
 PA (CADU-) CADUS PHARM CORP.
 XX
 PI Ciemoweki M, Duzic E;
 XX WPI; 2000-072337/06.
 DR P-PSDB; AAY53923.
 XX
 PT A new activator of G protein signalling used to treat disorders
 PT characterized by an aberrant AGS protein activity -
 XX
 PS Example 15, Page 143-144; 162pp; English.
 XX
 CC The present sequence encodes a homologue of a human AGS1 (activator
 CC of G protein signalling (AGS)) protein. The AGS cDNA sequence was
 CC isolated from a human liver cDNA library. The AGS protein exhibits
 CC homology to Ras-related G proteins, and contains alterations in
 CC conserved amino acids consistent with a deficiency in GTP hydrolysis
 CC activity. AGS stimulates G protein activity, G protein-coupled signal
 CC transduction and the pheromone response pathway in a receptor-independent
 CC manner. The AGS protein also shows G-gamma selectivity, as measured by
 CC growth assays in yeast expressing various mammalian G-gamma constructs,
 CC and tissue-specific expression, as measured by Northern blot analysis.
 CC The AGS protein can be used to screen for compounds that modulate
 CC cellular signal transduction. The protein is used to treat disorders
 CC characterized by an aberrant AGS protein activity or AGS nucleic acid
 CC expression.
 CC
 SO Sequence 837 BP; 177 A; 272 C; 233 G; 155 T; 0 other;
 SO
 Alignment Scores:
 Pred. No.: 1,46e-93 Length: 837
 Score: 893.00 Matches: 174
 Percent Similarity: 78.14% Conservative: 44
 Best Local Similarity: 62.37% Mismatches: 45
 Query Match: 61.59% Indels: 16
 DB: 21 Gaps: 4

US-09-709-103-1f1 (1-282) x AA236913 (1-837)

QY 5 AlaMetIleLysLysMetCysProSerAspSerGluLeuSerIleProAlaLysAsnGys 24
 Db 34 GCCATGATGAAGACTTTGTCCAGCGGAACGTCACGCTCAGTGTGCGCCGCAAAATCA 93
 QY 25 TyrArgMetValIleLeuGlySerSerLysValGlyLysThrAlaIleValSerArgPhe 44
 Db 94 TACCGCATGTGTGTCTGGTGTCTCTCGGGTGGCAAGACTCCATGTCCTTCGCTTC 153
 QY 45 LeuThrGlyArgPheGluAspAlaIleThrProThrIleGluAspPheIleArgLysPhe 64
 Db 154 CTCATGCGCGCTTTGAGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 213
 QY 65 TyrSerIleArgGlyGluValTyrGluLeuAspIleLeuAspThrSerGlyAsnIlePro 84
 Db 214 TACAACATCCCGCGGACGATGACGACGACGACGACGACGACGACGACGACGACGACGAC 273
 QY 85 PheProAlaMetArgArgLysLeuSerIleLeuThrGlyAspValIlePheIleLeuValPheSer 104
 Db 274 TTCCTGGCATGCGGAGGCTGTCTCTCAAGGGAGTGTCTTCATCTCGTGTTCAGC 333
 QY 105 LeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGlnGlnIleLeuAspThr 124
 Db 334 CTGATTAACCGGAGACTCTTCGATGATGAGTCAAGCCCTTCGAAAGCAGATCTGAGAGTC 393
 QY 125 LysSerCysLeuLysAsnLysThrLysGluAsnValAspValProLeuValIleCysGly 144

Db 394 AAGTCTCGCTGAAGAACAGACGAGGCGCGGAGCTGCTCATGTCTGTGC 453
 QY 145 AsnLysGlyAspArg---AspPheTyrArgGluValAspGluAlaIleGluGlnLeu 163
 Db 454 AACCAAGAACGACGAGCGAGCTGTGCGCGAGTGTCCACGACCGAGCGAGCTGTG 513
 QY 164 ValGlyAspAspProGlnArgCysAlaTyrPheGluIleSerAlaLysLysAsnSerSer 183
 Db 514 GTGTGCGGCGAC---GAGAACTGCGCTTACTTGAAGTGTGCGGCAAGAAAGAACCAAC 570
 QY 184 LeuAspGlnMetPheArgAlaLeuPheAlaMetAlaLysLeuProSerGluMetSerPro 203
 Db 571 GTGACGAGATGTTCTTACGCTCTTCAGCATGCGCAAGCTGCCACAGAGATGAGCCCC 630
 QY 204 AspLeuHisArgLysValSerValGlnTyrCysAspValLeuHisLysLys-----Ala 221
 Db 631 GCCCTGATGCAAGATCTTCGCGAGTACGGTGAAGCGCTTCCAGCCGCTTCTGCG 690
 QY 222 LeuArgAsnLysLysLeuLeuArgAlaGlySerGlyGlyGlyAspProGlyAsp 241
 Db 691 ATGCGCGCGCTCAAGAGATG-----GAC 714
 QY 242 AlaPheGlyTleValAlaProPheAlaArgArgProSerValHisSerAspLeuMetTyr 261
 Db 715 GCTATGAGCATGTGTGCGCTTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 774
 QY 262 ILAArgGluLysLysAspAlaGlySerGlnAlaLysAspLysGluArgGlyValIle 280
 Db 775 ATCAAGGCCAAGGTCTCTCGGGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 831

RESULT 10
 ABL92076
 ID ABL92076 standard, cDNA; 2832 BP.
 XX
 AC ABL92076;
 XX
 DT 30-MAY-2002 (first entry)
 XX
 DE Human Tumour Endothelial Marker polynucleotide SEQ ID NO 174.
 XX
 KW Human; mouse; rat; TEM; tumour endothelial marker; NEM; PEM; cytostatic;
 KW normal endothelial marker; pan-endothelial marker; immunostimulant;
 KW antiangiogenic; tumour; neovascularization; vasculatured tumour;
 KW polycystic kidney disease; diabetes; retinopathy; rheumatoid arthritis;
 KW psoriasis; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN MO200210217-A2.
 XX
 PD 07-FEB-2002.
 XX
 PF 01-AUG-2001; 2001WO-US24031.
 XX
 PR 02-AUG-2000; 2000US-222599P.
 PR 11-AUG-2000; 2000US-224360P.
 PR 11-APR-2001; 2001US-282850P.
 XX
 PA (UY00) UNITV JOHNS HOPKINS.
 XX
 PI St Croix B, Kinzler KW, Vogelstein B;
 XX
 DR WPI; 2002-291856/33.
 XX
 PT An isolated molecule comprising an antibody variable region which
 PT specifically binds to an extracellular domain of a tumor endothelial
 PT marker (TEM) protein, useful for inhibiting tumor growth -
 XX
 PS Disclosure; Page 119-120; 33pp; English.
 XX
 CC The invention relates to an isolated molecule comprising an antibody
 CC variable region which specifically binds to an extracellular domain of a

CC tumour endothelial marker (TEM) protein selected from ABB90732, ABB90740,
 CC ABB90749, ABB90750 and ABB90769. The antibodies which bind to TEM
 CC proteins have cytostatic, immunostimulant and antiangiogenic activity.
 CC They are useful for inhibiting tumour growth, neoangiogenesis in
 CC subjects bearing a vascularised tumour, polycystic kidney disease,
 CC diabetic retinopathy, rheumatoid arthritis and psoriasis. Human, mouse
 CC and rat TEM genes and the encoded proteins (ABL92075-ABL92141 and
 CC ABB90721-ABB90789) are disclosed, as are marker oligonucleotide
 CC sequences: tumour endothelial markers (TEM) ABL91996-ABL92041 and
 CC ABL92143-ABL92191; normal endothelial markers (NEM) ABL92042-ABL92074;
 CC and pan-endothelial markers (PEM) ABL91903-ABL91995.

XX Sequence 2832 BP; 602 A; 891 C; 727 G; 612 T; 0 other;

Alignment Scores:

Pred. No.: 8.3e-93 Length: 2832
 Score: 893.00 Matches: 174
 Percent Similarity: 78.14% Conservative: 44
 Best Local Similarity: 62.37% Mismatches: 45
 Query Match: 61.59% Indels: 16
 DB: 24 Gaps: 4

US-09-709-103-1F1 (1-282) x ABL92076 (1-2832)

Qy 5 AlaMetIleLysLysMetCysProSerAspSerGluLeuSerIleProAlaLysAsnCys 24
 Db 34 GCCATGATGAAGACTTTGTCCAGCGGAAGTGCAGCTCAGTGTGCCCGCCAAACTCA 93
 Qy 25 TyrArgMetValIleLeuGlySerSerLysValGlyLysThrAlaIleValSerArgPhe 44
 Db 94 TACCGCATGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGT 153
 Qy 45 LeuThrGlyArgPheGluAspAlaTyrThrProThrIleGluAspPheHisArgLysPhe 64
 Db 154 CTCAATGCCCGCTTTAGGACCCAGTACACACCCACCACCTCCAGGACTTCCACCGCTAAGGTA 213
 Qy 65 TyrSerIleArgGlyGluValTyrGlnLeuAspIleLeuAspThrSerGlyAsnHisPro 84
 Db 214 TACAACATCCCGCGGACATGTACACATCTGCATCTCGATACCTTGGCAACACCCCC 273
 Qy 85 PheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIleLeuValPheSer 104
 Db 274 TTCCCCCGCATGCCAGGCTGTCTCATCTCACAGGGATGTCTTCATCTCTGGTTCAGC 333
 Qy 105 LeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGlnGlnIleLeuAspThr 124
 Db 334 CTGGATAACCGGGAGTCTCTCGATGAGTCAAGCGCTTCAGAAGCAGATCCTGGAGTC 393
 Qy 125 LysSerCysLeuLysAsnLysThrLysGluAsnValAspValProLeuValIleCysGly 144
 Db 394 AAGTCCTGCTGAAGAACCAAGACCAAGAGCGCGGAGCTGCCCATGGTTCATCTGTGGC 453
 Qy 145 AsnLysGlyAspArg---AspPheTyrArgGluValAspGlnArgGluIleGluGlnLeu 163
 Db 454 AACAAAGAACGACCCAGCGGAGCTGTGCCGCGAGTGTGCCACCCAGCGGAGCTGTG 513
 Qy 164 ValGlyAspAspProGlnArgCysAlaTyrPheGluIleSerAlaLysLysAsnSerSer 183
 Db 514 GTGTCCGGCGAC---GAGAACTGCGCTTCTTCAGGCTGTCCGCAAGAAAGAACCAAC 570
 Qy 184 LeuAspGlnMetPheArgAlaLeuPheAlaMetAlaLysLeuProSerGluMetSerPro 203
 Db 571 GTGGACGAGATGTTCTTACGTCTCTTACGATGTGCCAAGCTGCCACACGAGATGAGCCC 630
 Qy 204 AspLeuHisArgLysValSerValGlnTyrCysAspValLeuHisLysLys-----Ala 221
 Db 631 GCCCTGTCATCGCAGATCTCCGTGAGTACGCTTCCAGCTTCCAGCCCGCTTCTTGC 690
 Qy 222 LeuArgAsnLysLysLeuLeuArgAlaGlySerGlyGlyGlyGlyAspProGlyAsp 241
 Db 691 ATGCCCGCGCTCAGGAGATG-----GAC 714
 Qy 242 AlaPheGlyIleValAlaProPheAlaArgProSerValHisSerAspLeuMetTyr 261

Db 715 GCCTATGCGCATGGTCTCGCCCTTCCCGCGCCGCGCCAGCTCAACAGTGACCTCAAGTAC 774
 Qy 262 IleArgGluLysAlaSerAlaGlySerGlnAlaLysAspLysGluArgCysValIle 280
 Db 775 ATCAAGGCCAAGGTCTCTCGGGAAGGCCAGCGCGTGAGAGGGGCAAGTGCACCATC 831

RESULT 11

ABL92087
 ID ABL92087 standard; cDNA; 2973 BP.

XX ABL92087;

XX 30-MAY-2002 (first entry)

XX Human Tumour Endothelial Marker polynucleotide SEQ ID NO 197.

XX Human; mouse; rat; TEM; tumour endothelial marker; NEM; PEM; cytostatic;
 KW normal endothelial marker; pan-endothelial marker; immunostimulant;
 KW antiangiogenic; tumour; neoangiogenesis; vascularised tumour;
 KW polycystic kidney disease; diabetes; retinopathy; rheumatoid arthritis;
 KW psoriasis; gene; ss.

XX Homo sapiens.

XX WO200010217-A2.

XX 07-FEB-2002.

XX 01-AUG-2001; 2001WO-US24031.

XX 02-AUG-2000; 2000US-222599P.

XX 11-AUG-2000; 2000US-224360P.

XX 11-APR-2001; 2001US-282850P.

XX (UWJO) UNIV JOHNS HOPKINS.

XX St Croix B, Kinzler KW, Vogelstein B;

XX WPI; 2002-291856/33.

XX P-PSDB; ABB90733.

XX An isolated molecule comprising an antibody variable region which
 PT specifically binds to an extracellular domain of a tumor endothelial
 PT marker (TEM) protein, useful for inhibiting tumor growth -

XX Claim 65; Page 152-153; 33pp; English.

XX The invention relates to an isolated molecule comprising an antibody
 CC variable region which specifically binds to an extracellular domain of a
 CC tumour endothelial marker (TEM) protein selected from ABB90732, ABB90740,
 CC ABB90749, ABB90750 and ABB90769. The antibodies which bind to TEM
 CC proteins have cytostatic, immunostimulant and antiangiogenic activity.
 CC They are useful for inhibiting tumour growth, neoangiogenesis in
 CC subjects bearing a vascularised tumour, polycystic kidney disease,
 CC diabetic retinopathy, rheumatoid arthritis and psoriasis. Human, mouse
 CC and rat TEM genes and the encoded proteins (ABL92075-ABL92141 and
 CC ABB90721-ABB90789) are disclosed, as are marker oligonucleotide
 CC sequences: tumour endothelial markers (TEM) ABL91996-ABL92041 and
 CC ABL92143-ABL92191; normal endothelial markers (NEM) ABL92042-ABL92074;
 CC and pan-endothelial markers (PEM) ABL91903-ABL91995.

XX Sequence 2973 BP; 596 A; 967 C; 787 G; 623 T; 0 other;

Alignment Scores:

Pred. No.: 8.89e-93 Length: 2973
 Score: 893.00 Matches: 174
 Percent Similarity: 78.14% Conservative: 44
 Best Local Similarity: 62.37% Mismatches: 45
 Query Match: 61.59% Indels: 16
 DB: 24 Gaps: 4

US-09-709-103-1F1 (1-282) x ABL92087 (1-2973)

QY 5 AlameTlleYslySmetCysProSerAspSerGluLeuSerilleProAlaIysAncys 24
 DB 205 GCCATGATGAAGACTTGTTCACAGCGGAACTGACGCTCAGTGTGCGCCCAAAAACCTCA 264
 QY 25 TyrArgMetValIleLeuLeuIleSerIleValGlyValThAlaIleValSerArgPhe 44
 DB 265 TACCGATGATGATGCTGTGCTGCTCTCGGGTGGCGCAAGAGCTCCATCTGCTGCTTC 324
 QY 45 LeuThGlyArgPheGluAspAlaTyrTrnProThrIleGluAspPheHisArgIysPhe 64
 DB 325 CTCGAATGCGCGCTTTGAGAGACAGTACACACCCATGAGAGACTTCCACCGTAAGGTA 384
 QY 65 TyrSerIleArgGlyGluValIleTyrGlnLeuAspIleLeuAspThrSerGlyAsnHisPro 84
 DB 385 TACCAATCCGCGGACATGTATACAGCTGACATCTGATACCTCTGGAACCAACCC 444
 QY 85 PheProAlaMetArgArgIleuSerIleLeuThGlyAspValPheIleLeuValPheSer 104
 DB 445 TTCCTCCGATGCGAGGCTGTCTCCATCTCAACAGGGAGATGCTTCACTCTGCTGTACG 504
 QY 105 LeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGlnIleLeuAspThr 124
 DB 505 CTGGATTAACCGGAGAGCTTTCATAGAGTCAAGCCCTTCAAGAGCAATCTTGAAGGTC 564
 QY 125 IysSerCysLeuIlysAsnIysThrIysGluAsnValAspValProLeuValIleCysGly 144
 DB 565 AAGTCTGCTGTAAGAACAGACCAAGAGAGGGCGGAGCTGCCATGCTACTCTGTGGC 624
 QY 145 AsnIysGlyAspArg---AspPheTyrArgGluValAspGlnArgIleGluIleu 163
 DB 625 AACCAAGAACGACACCGGACGCTGTGCTCCGACAGTCCACACCGAGCGCAGCTGTCTG 684
 QY 164 ValGlyAspAspProGlnArgCysAlaTyrPheGluIleSerAlaIysIysAsnSer 183
 DB 685 GTCTCGGGGAC---GAGAACTGCGCTTCTTCAAGTGTGCGCAAGAGAACACCAAC 741
 QY 184 LeuAspGlnMetPheArgIleLeuPheAlaMetAlaIysLeuProSerGluMetSerPro 203
 DB 742 GTGAGCAGAGATGTTCTACGCTGCTTCAAGATGGCCAGCTGCCACAGAGATGAGCC 801
 QY 204 AspLeuHisArgIysValSerValGlnTyrCysAspValLeuHisIysLys---Ala 221
 DB 802 GCCCTGATCGCAAGATCTCCGTGCAAGTACGCTTCCAGCCAGCGCTTCTGTC 861
 QY 222 LeuArgAsnIysLeuLeuArgIleArgIleSerGlyGlyGlyIleAspProIlyAsp 241
 DB 862 ATGCGCGCGCTCAAGGAGATG-----GAC 885
 QY 242 AlaPheGlyIleValAlaProPheAlaArgArgProSerValHisSerAspLeuMetTyr 261
 DB 886 GCGTATGCGATGCTGCTCGCTTCCGCGCGCGCCAGGCTCAACAGAGCTTCAAGTAC 945
 QY 262 IleArgGluIysAlaSerAlaGlySerGlnAlaIysAspIysGluArgCysValIle 280
 DB 946 ATCAAGGCGCAAGTCTTCTCGGAGAGCGCAGCGCTGAGAGGAGCAAGTGCACCATC 1002
 RESULT 12
 ABK71563
 ID ABK71563 standard; cDNA, 3427 BP.
 XX ABK71563;
 AC 30-JUL-2002 (first entry)
 DT
 XX
 DE Human dithp polynucleotide #29.
 XX
 KW Human; dithp; diagnostic and therapeutic polynucleotide; gene; ss; bone;
 KW cell proliferative disorder; cancer; tumour; autoimmune disorder; brain;
 KW inflammatory disorder; viral infection; bacterial infection; seizure;
 KW fungal infection; parasitic infections; developmental disorder; breast;
 KW endocrine disorder; metabolic disorder; neurological disorder; cervix;
 KW gastrointestinal disorder; transport disorder; gene therapy; kidney;

KW adrenal gland; bone marrow; lung; ovary; pancreas; prostate; spleen;
 KW skin; testis; thymus.
 XX
 OS Homo sapiens.
 XX
 FN W0200220754-A2.
 XX
 PD 14-MAR-2002.
 XX
 PF 29-AUG-2001; 2001MO-US27127.
 XX
 PR 05-SEP-2000; 2000US-229747P.
 PR 05-SEP-2000; 2000US-229748P.
 PR 05-SEP-2000; 2000US-229749P.
 PR 05-SEP-2000; 2000US-229750P.
 PR 05-SEP-2000; 2000US-229751P.
 PR 05-SEP-2000; 2000US-230583P.
 PR 06-SEP-2000; 2000US-230584P.
 PR 06-SEP-2000; 2000US-230514P.
 PR 06-SEP-2000; 2000US-230515P.
 PR 06-SEP-2000; 2000US-230517P.
 PR 06-SEP-2000; 2000US-230518P.
 PR 06-SEP-2000; 2000US-230519P.
 PR 06-SEP-2000; 2000US-230597P.
 PR 06-SEP-2000; 2000US-230599P.
 PR 06-SEP-2000; 2000US-230598P.
 PR 06-SEP-2000; 2000US-230599P.
 PR 06-SEP-2000; 2000US-230610P.
 PR 06-SEP-2000; 2000US-230656P.
 PR 06-SEP-2000; 2000US-230988P.
 PR 07-SEP-2000; 2000US-230951P.
 PR 07-SEP-2000; 2000US-231163P.
 PR 07-SEP-2000; 2000US-231167P.
 XX
 PA (INCR-) INCTE GENOMICS INC.
 XX
 PI Stuart J, Lincoln SE, Altus CM, Dufour GE, Chalup MS, Hillman JL;
 PI Jones AU, Yu JY, Wright RJ, Gierzen D, Liu TF, Yap PE, Dahl CR;
 PI Miyama MG, Bradley DL, Rohatgi SD, Harris B, Roseberry AM;
 PI Gerstein EH, Peralta CH, David MH, Panzer SR, Florez V, Datto A;
 PI Marwaha R, Chen AV, Chang SC, Au AP, Inman RR;
 XX
 DR WPI; 2002-383054/41.
 DR P-PSDB; ABG59971.
 XX
 PT An isolated polynucleotide useful in diagnostics and therapeutics -
 XX
 PS Claim 1; Page 419-420; 686pp; English.
 XX
 CC The invention relates to human diagnostic and therapeutic (dithp)
 CC polynucleotides and their associated polypeptides (Dithp polypeptides).
 CC The sequences of the invention are used in the treatment and diagnosis of
 CC cell proliferative disorders (e.g. atherosclerosis, cirrhosis), cancers
 CC (e.g. tumours of the adrenal gland, bone, bone marrow, brain, breast,
 CC cervix, kidney, lung, ovary, pancreas, prostate, skin, spleen, testis or
 CC thymus), autoimmune/inflammatory disorders (e.g. asthma, bronchitis,
 CC psoriasis, osteoporosis), viral infections, bacterial infections, fungal
 CC infections, parasitic infections, developmental disorders (e.g. anaemia,
 CC epilepsy), seizure disorders (e.g. cerebral palsy, spina bifida),
 CC endocrine disorders (e.g. thrombosis, aneurysm), metabolic disorders
 CC (e.g. obesity, diabetes), neurological disorders (e.g. stroke,
 CC amyotrophic lateral sclerosis, multiple sclerosis), gastrointestinal
 CC disorders (e.g. ulcerative colitis, lysinuria) and transport disorders
 CC (e.g. myotonic dystrophy, catatonias, peripheral neuropathy). Sequences
 CC ABK71535-ABK71809 represent human dithp polynucleotides of the invention.
 XX
 SQ Sequence 3427 BP; 682 A; 1145 C; 935 G; 665 T; 0 other;
 Alignment Scores:
 Pred. No.: 1 09e-92 Length: 3427
 Score: 893.00 Matches: 174
 Percent Similarity: 78.14% Conservative: 44
 Best Local Similarity: 62.37% Mismatches: 45

Query Match:	61.5%	Indels:	16
DB:	24	Gaps:	4
US-09-709-103-1F1 (1-282) x ABK71563 (1-3427)			
Qy	5	AlaMetIleLysLysMetCysProSerAspSerGluLeuSerIleProAlaLysAsnCys	24
Db	641	GCCATGATGAAGACTTTTGTCCACGCGGAATGCACCGCTCAGTGTGCCGCCAAAACATCA	700
Qy	25	TyrArgMetValIleLeuGlySerSerLysValGlyLysThrAlaIleValSerArgPhe	44
Db	701	TACCGCATGTGGTGTGGGTGCTTTCGGTGGCAAGAGCTCCATCGTGTCTCGCTTC	760
Qy	45	LeuThrGlyArgPheGluAspAlaTyrThrProThrIleGluAspPheHisArgLysPhe	64
Db	761	CTCAATGCCCGCCTTGAGGACAGTACACACCCACCATCGAGACTTCACCGTAAGGTA	820
Qy	65	TyrSerIleArgGlyGluValTyrGlnLeuAspIleLeuAspThrSerGlyAsnHisPro	84
Db	821	TACAACATCCGCGCAGATGTACAGCTGCATCTCGATACCTCTCGCAACACACCCC	880
Qy	85	PheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIleLeuValPheSer	104
Db	881	TTCCCCGCCATGCGCAGCGTCCATCCTCACAGGGAGTCTTTCATCTCGTGTTCAGC	940
Qy	105	LeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGlnGlnIleLeuAspThr	124
Db	941	CTGGATTAACCGGGAGTCTTCGATAGGTCAAGCGCTTCAGAAGCAGATCCTGGAGGTC	1000
Qy	125	LysSerCysLeuLysAsnLysThrLysGluAsnValAspValProLeuValIleCysGly	144
Db	1001	AAGTCTCTGCTGAAGAACAAAGACCAAGGAGCGCGGAGCTGCCATCTGTCTCTGTGC	1060
Qy	145	AsnLysGlyAspArg---AspPheTyrArgGluValAspGlnArgGluIleGluGlnLeu	163
Db	1061	AACAAGAACACACACGCGAGCTGTGCGCGCCAGGTGCCACACCAGCGCCGAGCTGCTG	1120
Qy	164	ValGlyAspAspProGlnArgCysAlaTyrPheGluIleSerAlaLysLysAsnSerSer	183
Db	1121	GTGTGCGGCGAC---GAGAACTCGCTACTTCGAGGTGTGCGCCAAAGAACACCAAC	1177
Qy	184	LeuAspGlnMetPheArgAlaLeuPheAlaMetAlaLysLeuProSerGluMetSerPro	203
Db	1178	GTGCACGAGATGTTCTACGTGCTCTTCAGCATGCCAAGCTGCCACACAGATGAGCCCC	1237
Qy	204	AspLeuHisArgLysValSerValGlnTyrCysAspValLeuHisLysLys	221
Db	1238	GCCCTGCATGCCAAGATCTCCGTGCAGTACGGTGACGCTTCACCCCGAGCCCTTCCTG	1297
Qy	222	LeuArgAsnLysLysLeuLeuArgAlaGlySerGlyGlyGlyGlyAspProGlyAsp	241
Db	1298	ATGCCCGCGCTCAGGAGATG-----GAC	1321
Qy	242	AlaPheGlyIleValAlaProPheAlaArgProSerValHisSerAspLeuMetTyr	261
Db	1322	GCCTATGCGATGTCTCGCCCTTCGCCCGCGCCAGCTCAACAGTGCACCTCAAGTAC	1381
Qy	262	IleArgGluLysAlaSerAlaGlySerGlnAlaLysAspLysGluArgCysValIle	280
Db	1382	ATCAAGGCCAAGTCTCTCGGGAAGCGCCGCTGAGAGGACCAAGTGCACCATC	1438
RESULT 13			
ABL92134			
ID	ABL92134 standard; cDNA; 3020 BP.		
XX	AC		
XX	AC		
XX	AC		
DT	30-MAY-2002 (first entry)		
XX	Mouse Tumour Endothelial Marker polynucleotide SEQ ID NO 292.		
DE	Human; mouse; rat; TEM; tumour endothelial marker; NEM; PEM; cytotstatic;		
XX	normal endothelial marker; pan-endothelial marker; immunostimulant;		
KW			

Db 613 TTCCCTGCGATGCGCGGCTCTCCATCTCAAGAGATGCTTCTTCCTGCTGCTTACG 672
 Qy 105 LeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGlnGlnLeuAspThr 124
 Db 673 CTGATATGCCCCGGAGTCTCTTGTAGAGTCAAGCGCTCCAGAAACAGATCTTGAGAGTC 732
 Qy 125 LysSerCysLeuValAsnLeuThrValGluAsnValAspValProLeuValIleCysGly 144
 Db 733 AAGTCTGCTCTGAAGAATTAACCAAGAGAGCAGCAGAGCTGCCATGATGATCTGGGG 792
 Qy 145 AsnLysGlyAspArg---AspPheTyrArgGluValAspGlnArgGlnIleGluGlnLeu 163
 Db 793 AACAAAGATGACCAAGAGAGCTGTGCGCCAGCTCCCTGCTGAGGCTGAGCTGCTG 852
 Qy 164 ValGlyAspAspProGlnArgCysAlaTyrPheGlnIleSerAlaIleLysLysAsnSerSer 183
 Db 853 GTGCTGCTGAT---GAAACTGCGCTTATTTCAGAGTGTACAGCAAGAAACACTAAT 909
 Qy 184 LeuAspGlnMetPheArgAlaLeuPheAlaMetAlaLysLeuProSerGlnMetSerPro 203
 Db 910 CTGACAGAGATGTTCTTATGCTGCTTACAGATGCGCAAGCTGCCCCATGAGATGAGCCCT 969
 Qy 204 AspLeuHisArgLysValSerValGlnTyrCysAspValLeuHisLysLysAlaLeuArg 223
 Db 970 GCACTGACCATTAAGATCTCCGTCAGATAGCGGAGTCTTTTCACCCCCGCTTCTGC 1029
 Qy 224 AsnLysLysLeuLeuArgAlaGlySerGlyGlyGlyGlyAspProGlnAspAlaPhe 243
 Db 1030 ATGCGTGGCACTAAGTGTGCGAGT-----GCCTAAT 1059
 Qy 244 GlyIleValAlaProPheAlaArgArgProSerValHisSerAspLeuMetLysLeuArg 263
 Db 1060 GCGATGCTTCAACCTTTGCGCCAGCGCCCAAGTGTCAACAGTACTCAAGATCAATCAAG 1119
 Qy 264 GluLysAlaSerAlaGlySerGlnAlaLysAspLysGluArgCysValIle 280
 Db 1120 GCCAAGTCTTACGAGGAGGCGCCAGCCAGAGAGGAGAACAGTGTAGCATC 1170
 RESULT 14
 AAS90571
 ID AAS90571 standard; cDNA, 951 BP.
 XX AAS90571;
 AC
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE DNA encoding novel human diagnostic protein #26375.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US08631.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 XX
 PR 23-AUG-2000; 2000US-0649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;
 XX
 DR WPI, 2001-639362/73.
 DR P-PsDB; ABG26384.
 XX
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.

XX
 PS Claim 1; SEQ ID No 26375; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridization probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC and to produce other types of data or other traits to assess biodiversity
 CC amino acid sequences. AAS64197-AAS94564 represent novel human
 CC diagnostic coding sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX

SQ Sequence 951 BP; 202 A; 291 C; 293 G; 165 T; 0 other:

Alignment Scores:
 Pred. No.: 4.06e-73 Length: 951
 Score: 717.00 Matches: 149
 Percent Similarity: 69.78% Conservative: 38
 Best Local Similarity: 55.60% Mismatches: 45
 Query Match: 49.45% Indels: 36
 DB: 23 Gaps: 7

US-09-709-103-1f1 (1-282) x AAS90571 (1-951)

Qy 26 ArgMetValIleLeuGlySerSerLysValGlyLysThrAlaIleValSerArgPheLeu 45
 Db 142 CGCATGGTGGTGTGCTGGGTGCTCGGGTGGCAAGCTCCATCGTGTCTTCTTCTC 201
 Qy 46 ThrGlyArgPheGluAspAlaTyrThrProThrIleGluAspPheHisArgLysPheTyr 65
 Db 202 AATGCCCGCTTTGAGGACCACTACACACCACCATCGAGGACTTCCACCGTAAAGTATAC 261
 Qy 66 SerIleArgGlyValValTyrGlnLeuAspIleLeuAspThrSerGlyAsnHisProPhe 85
 Db 262 AACATCGCGCGGCGACATGTACAGCTGACATCTGTGATCCTTGGCAACCAACCCCTTC 321
 Qy 86 ProAlaMetArgArgLysSerIleLeuThrGlyAspVal---PheIleLeuValPheSer 104
 Db 322 CCGCGCATGGCGAGCGTGTCTCATCTTCACAGGTGAGCCCACTGTGCTTGGGCTGGGCGC 381
 Qy 105 LeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGln----- 119
 Db 382 -----GCCAGGCGCCAGGCGCATGGTGGCGAGTGTG 411
 Qy 120 -----GlnIleLeuAspThrLysSerCysLeuLysAsnLysThrLysGluAsn 135
 Db 412 CTGGGCACTTAAGACAGATCTGAGGTCAAGTCTGCTGAAGAACCAAGCAAGAGGCGG 471
 Qy 136 ValAspValProLeuValIleCysGlyAsnLysGlyAspArg---AspPheTyrArgGln 154
 Db 472 GCGGAGCTGCCATGTGATCTGTGGCAACAAAGACCAACCGGACGCTGTGCGCCGCG 531
 Qy 155 ValAspGlnArgGluIleGluGlnLeuValGlyAspAspProGlnArgCysAlaTyrPhe 174
 Db 532 GTGCGCCACCAACGAGCGCAGCTGTGCTGGTGGCGGAC---GAGAACTGCGCCCTAATTG 588
 Qy 175 GluIleSerAlaLysLysAsnSerSerLeuAspGlnMetPheArgAlaLeuPheAlaMet 194
 Db 589 GAGGTGTGCGCCAAAGAAACCAACCAAGTGTGACGATGTCTTACCTGCTTTCAGCATG 648

QY 195 AlaLysLeuProSerGluMetSerProAspLeuHisArgLysValSerValGlnTyrCys 214
 DB 649 GCCAAGCTGCCACAGAGATGAGCCCGCTGCATCGCAAGATCTCCGTCAGTACGGT 708
 QY 215 AspValLeuHisLysLys-----AlaLeuArgAsnLysLysLeuLeuArgAlaGlySer 232
 DB 709 GAGCGCTTCCACCCAGGCGCTTCTGCATCGCGCGCTCAAGGAGATG----- 756
 QY 233 GlyGlyGlyGlyGlyAspProGlyValAlaPheGlyLeuValAlaProPheAlaArgArg 252
 DB 757 -----GAGCGCTATGGCATGGTCTCGCCCTTCGCCCTTCGCCCGCGCG 792
 QY 253 ProSerValHisSerAspMetTyrLeuArgGlyLysAlaSerAlaGlySerGlnAla 272
 DB 793 CCCAGCGTCAACAGTACCTCAAGTACATCAAGCCCAAGGTCTTCGGGAAGGCCAGGCC 852
 QY 273 LysAspLysGluArgCysValile 280
 DB 853 CGTGAGGGGACAAGTGCACCATC 876

RESULT 15
 ABL07789
 ID ABL07789 standard; cDNA; 1305 BP.
 XX AC
 XX ABL07789;
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 17849.
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical; gene; ss.
 OS Drosophila melanogaster.
 XX
 XX WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US09231.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 XX (PEKE) PE CORP NV.
 PA
 XX Venter JC, Adams M, Li PWD, Myers EW;
 PI WPI; 2001-656860/75.
 DR P-PSDB; ABB63686.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX
 PS Claim 1; SEQ ID NO 17849; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABLL16176-ABLL30511), expressed DNA
 CC sequences (ABLL01840-ABLL16175) and the encoded proteins
 CC (ABBS7737-ABBS72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 1305 BP; 313 A; 390 C; 336 G; 266 T; 0 other;

Alignment Scores: 4.04e-61 Length: 1305
 Pred. No.:

Score: 615.00 Matches: 134
 Percent Similarity: 64.42% Conservative: 38
 Best Local Similarity: 50.19% Mismatches: 69
 Query Match: 42.41% Indels: 26
 DB: 23 Gaps: 6

US-09-709-103-1F1 (1-282) x ABL07789 (1-1305)

QY 18 SerIlePro---AlaLysAsnCysTyrArgMetValIleLeuGlySerSerLysValGly 36
 DB 475 AGCCTGCCATCGGCGCAAACTGCTATCGATGGTTATGTCGGCTCATCACGCGCGCG 534
 QY 37 LysThrAlaIleValSerArgPheLeuThrGlyArgPheGluAspAlaTyrThrProThr 56
 DB 535 AAGTCATCGATTGTGGCAGCTTCTCTGGGCAATCGTTTCGAGGAGGCTTACACGCGGACC 594
 QY 57 IleGluAspPheHisArgLysPheTyrSerIleArgGlyGluValTyrGlnLeuAspIle 76
 DB 595 ATCGAGGAGTTCACCGCAATTTGATCGCATCGGAATGAAGTCTTTCAATTGGATATT 654
 QY 77 LeuAspThrSerGlyAsnHisProPheProAlaMetArgArgLeuSerIleLeuThrGly 96
 DB 655 TTGGATACCTCTGGCTATCATCCGTTTCCGCAATGCGACGTTTATCATTTCTAACTGGG 714
 QY 97 AspValPheIleValPheSerLeuAspAsnArgAspSerPheGluGluValGlnArg 116
 DB 715 GATCTCTTCATCTCTGCTTCAGCATGGATTCCCGCGAGTCTTTCGAGGAGGTGGTTCGC 774
 QY 117 LeuArgGlnGlnIleLeuAspThrLysSerCysLeuLysAsn-----Lys 131
 DB 775 CTGCGGGAGACATCTTGGAGACCAAGTGGGCTGCATAAATCCCGGCTCCGGGTTTAAAG 834
 QY 132 ThrLysGluAsnValAspValProLeuValIleCysGlyAsnLysGlyAspArgPhe 151
 DB 835 AAGAGAGTCTTCCAAAGATACCCATGATATTTGGCGGGAATAAATGATCGAGACTTT 894
 QY 152 TyrArgGluValAspGlnArgGluIleGluGlnLeuValGlyAspAspProGlnArgCys 171
 DB 895 ---AAACTGTTCAAGTGACGAGGATGGGCTACATCGTGGCCAGGACAACTGCTGC 951
 QY 172 AlaTyrPheGluIleSerAlaLysLysAsnSerSerLeuAspGlnMetPheArgAlaLeu 191
 DB 952 ACCTTTGTGGAGTGTCTCGCTCGTCAGAAATTACCGCATCGATGACCTATTCCACTCGCTG 1011
 QY 192 PheAlaMetAlaLysLeuProSerGluMetSerProAspLeuHisArgLys---ValSer 210
 DB 1012 TTTACGCTCTCCAATCTCGCGTGGAGATGACCCCAATCATCATCTCGTGGTGTCTCC 1071
 QY 211 ValGlnTyrCysAspValLeuHisLysLysAlaLeuArgAsnLysLysLeuLeuArgAla 230
 DB 1072 GTTTTGGGCGGCC-----TCGCCACTCCCAACCCACAC 1104
 QY 231 GlySerGlyGlyGlyGlyAspPro-----GlyAsp 241
 DB 1105 GGATCTCGGTGGCGGGGACCAAGAAGATGCCCTGTCCATCAAGAGAGATTACCGCAT 1164
 QY 242 AlaphedyleValAlaProPheAlaArgArgProSerValHisSerAspLeuMetTyr 261
 DB 1165 GCCTGCGGTGTGTGACCGCGAAGCGGAGAGACCCGAGCATCCGACCGATCTCAATCTG 1224
 QY 262 IleArgGluLysAlaSerAla 268
 DB 1225 ATGAGATCCAGAGCGATGGCG 1245

Search completed: December 30, 2002, 14:45:52
 Job time : 366 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: December 27, 2002, 22:15:03 ; Search time 353 Seconds
(without alignments)
1792.667 Million cell updates/sec

Title: US-09-709-103-2
Sequence: 1 MFLAMIKMKCPSELISIP.....IREKASAGSQAKDERCVIS 281

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODE=frame-p2n.model -DEV=xlp
-O=/cgn2.1/USPTO.spool/US09709103/rnat.20122002.162405.797/app-query.fasta.1.455
-DB=Geneseq.101002 -OPMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=2000000000
-USER=US09709103.ecgn.1.1.79.grnat.20122002.162405.797 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-MARN TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
N_Geneseq.101002.*
1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the total being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query	Length	DB ID	Description
1	1449	100.0	846	AA236892	CDNA encoding an a
2	1449	100.0	1740	AA236914	DNA encoding an a
3	1449	100.0	1801	AA236893	CDNA encoding an a
4	1449	100.0	1841	AA236922	Human kd312 polype
5	1441	99.4	1776	AAA49177	CDNA encoding huma
6	1399.5	96.6	1689	AA233024	Rat kd312 polypept
7	1387	95.7	3986	AA233023	Human kd312 genom
8	1343	92.7	3079	AA233025	Rat kd312 genomic
9	893	61.6	837	AA236913	DNA encoding a hom
10	893	61.6	2832	ABL92076	Human Tumour Endot
11	893	61.6	2973	ABL92087	Human Tumour Endot
12	893	61.6	3427	ABK71563	Human dltph polynu
13	867	59.8	3020	ABL92134	Mouse Tumour Endot
14	717	49.5	951	AA590571	DNA encoding novel
15	615	42.4	1305	ABL07789	Drosophila melanog
16	494	34.1	624	ABQ28318	Oligonucleotide fo
17	494	34.1	624	ABQ28319	Oligonucleotide fo
18	459	31.7	624	ABQ28320	Oligonucleotide fo
19	459	31.7	624	ABQ28321	Oligonucleotide fo
20	370	25.5	368	ABA51382	Human breast cell
21	370	25.5	368	ABA59388	Human foetal liver
22	370	25.5	368	ABA36323	Probe #14789 for g
23	370	25.5	368	AAK17661	Human brain expres
24	370	25.5	368	AAK43477	Human bone marrow
25	370	25.5	368	AAI24259	Probe #14192 for g
26	370	25.5	368	AAI49541	Probe #18227 used
27	370	25.5	368	AAI09818	Probe #9809 used t
28	370	25.5	368	ABSI7600	Drosophila melanog
29	359.5	24.8	4543	ABL07788	Human polynucleoti
30	334	23.1	1087	AAI60838	Human polynucleoti
31	334	23.1	1108	AAI59052	Human polynucleoti
32	334	23.1	1167	AAI57605	Human intracellular
33	330.5	22.8	960	ABL22541	Drosophila melanog
34	322	22.2	3061	ABQ72648	Human MDT encodin
35	307.5	21.2	702	ABL14159	Drosophila melanog
36	305	21.0	1144	AAH90075	Human bone marrow
37	305	21.0	4933	AAK81235	Human immune/haema
38	305	21.0	4934	AAK81234	Human immune/haema
39	304	21.0	1249	ABQ72525	Human MDT encodin
40	302	20.8	558	AAQ03212	RAP2 Gene encoding
41	302	20.8	3300	AAI58569	Human polynucleoti
42	301	20.8	2205	ABL50199	Human Rapla, c-Raf
43	301	20.8	2205	ABL50231	Human Rapla, c-Raf
44	301	20.8	2277	ABL50202	Human Rapla, c-Raf
45	301	20.8	2277	ABL50234	Human Rapla, c-Raf

ALIGNMENTS

RESULT 1
AA236892
AA236892 standard; cDNA; 846 BP.

AC AA236892;

DT 13-MAR-2000 (first entry)

DE CDNA encoding an activator of G protein signalling (AGS) protein.

KW Activator of G protein signalling; AGS; ras-related G protein;

KW GTP hydrolysis; G protein activity; pheromone response pathway;

KW G protein-coupled signal transduction; G-gamma selectivity;

KW cellular signal transduction; ss.

OS Homo sapiens.
XX Location/Qualifiers
FH Key

```

FT CDS 1..846
FT /*tag= a
FT /product= "activator of G protein signalling (AGS)
FT protein"
Mo9958670-A1.
XX 18-NOV-1999.
XX
XX 07-MAY-1999; 99WO-US10151.
XX
XX 08-MAY-1998; 98US-0084842.
XX 07-OCT-1998; 98US-0103355.
XX
XX (CADU-) CADUS PHARM CORP.
XX
XX Cismowski M, Duzic E;
XX
XX WPI: 2000-072337/06.
XX P-PSDB; AA553921.
XX
XX A new activator of G protein signalling used to treat disorders
XX characterized by an aberrant AGS protein activity -
XX
XX Claim 3; Fig 3A; 162pp; English.
XX
XX The present sequence encodes an activator of G protein signalling (AGS)
XX protein. The cDNA sequence was isolated from a human liver cDNA
XX library. The AGS protein exhibits homology to ras-related G proteins,
XX and contains alterations in conserved amino acids consistent with a
XX deficiency in GTP hydrolysis activity. AGS stimulates G protein
XX activity, G protein-coupled signal transduction and the pheromone
XX response pathway in a receptor-independent manner. The AGS protein
XX also shows G-gamma selectivity, as measured by growth assays in
XX yeast expressing various mammalian G-gamma constructs, and
XX tissue-specific expression, as measured by Northern blot analysis.
XX The AGS protein can be used to screen for compounds that modulate
XX cellular signal transduction. The protein is used to treat disorders
XX characterized by an aberrant AGS protein activity or AGS nucleic acid
XX expression.
XX
XX Sequence 846 BP; 176 A; 288 C; 252 G; 130 T; 0 other;
XX
XX Alignment Scores:
XX Pred. No.: 6.91e-158 Length: 846
XX Score: 1449.00 Matches: 281
XX Percent Similarity: 100.00% Conservative: 0
XX Best local Similarity: 100.00% Mismatches: 0
XX Query Match: 100.00% Indels: 0
XX DB: 21 Gaps: 0
XX
XX US-09-709-103-2 (1-281) x AA236892 (1-846)
XX
XX QY 1 MetLysLeuAlaAlaMetIleLysLysMetCysProSerAspSerGluLeuSerIlePro 20
XX |||||||
XX DB 1 ATGAATCTGGCCGCGATATCAAGAAAGTGTGCCGAGCGACTGAGTATCCG 60
XX
XX QY 21 AlaLysAsnGlyTyrArgMetValIleLeuGlySerSerIysValGlyLysThrAlaIle 40
XX |||||||
XX DB 61 GCCAAGAACTGCTATCGCATGTCATCCCTCGTCGTCGCAAGGTTGGGCAAGAGCCATC 120
XX
XX QY 41 ValSerArgPheLeuThrGlyArgPheGluAspAlaTyrThrProThrIleGluAspPhe 60
XX |||||||
XX DB 121 GGTGCGCGCTTCACCGCGCCCTTCGAGACCCCTACACGCTTACATCGAGACTTC 180
XX
XX QY HisArgLysPheTyrSerIleArgGlyValTyrGlnLeuAspIleLeuAspThrSer 80
XX |||||||
XX DB 181 CACCGCAAGTCTACTCCATCCGCGGAGGTTACCAAGCTCGACATCTCGACAGCTCC 240
XX
XX QY 81 GlyAsnHisProPheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIle 100
XX |||||||
XX DB 241 GGCAACGACCGCTTCCCGGCATCGGGCGCTCTCCATCTCCACAGAGAGAGGTTTTCATC 300

```

```

QY 101 LeuValIPheSerLeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGlnGln 120
XX |||||||
XX DB 301 CTGGTGTTCAGTCTGGCAACACCGGACTCTCTCGAGAGAGTGCAGCGGCTCAGCAGCAG 360
XX
XX QY 121 IleLeuAspThrLysSerCysLeuLysAsnLysThrLysGluAsnValAspValProLeu 140
XX |||||||
XX DB 361 ATCTCGACACCAAGTGTCTCCAGAACCAAAACCAAGAGAACTGGACGCGCCCTG 420
XX
XX QY 141 ValIleCysGlyAsnLysGlyAspArgAspPheTyrArgGluValAspGlnArgGluIle 160
XX |||||||
XX DB 421 GTGATCTGGCGCAACAAAGGTGACCGGACTTCTACCGGAGGTGACCAAGCCGAGATC 480
XX
XX QY 161 GluGlnLeuValGlyAspAspProGlnArgCysAlaTyrPheGluIleSerAlaLysLys 180
XX |||||||
XX DB 481 GAGCAGCTGTGGGCGGACGCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
XX
XX QY 181 AsnSerSerLeuAspGlnMetPheArgAlaLeuPheAlaMetAlaLysLeuProSerGlu 200
XX |||||||
XX DB 541 AACGACAGCTGAGACCAAGATGTTCCGGCGCTTTCGCAATGGCCAAAGCTGCCAGCGAG 600
XX
XX QY 201 MetSerProAspLeuHisArgLysValSerValGlnTyrCysAspValLeuHisLysLys 220
XX |||||||
XX DB 601 ATGAGCCCAAGACCTGGACCGGCAAGGTCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
XX
XX QY 221 AlaLeuArgAsnLysLysLeuLeuArgAlaGlySerGlyGlyGlyGlyGlyGlyGlyGly 240
XX |||||||
XX DB 661 GCCTGTGGGAAACAAAGCTGTGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 720
XX
XX QY 241 AspAlaPheGlyIleValAlaProPheAlaArgArgProSerValHisSerAspLeuMet 260
XX |||||||
XX DB 721 GAGGCTTTGGATCGGGGACCTTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 780
XX
XX QY 261 TyrIleArgLysLysLysAlaSerAlaGlySerGlnAlaLysAspLysGluArgCysValIle 280
XX |||||||
XX DB 781 TATATCCGCGCAAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 840
XX
XX QY 281 Ser 281
XX |||
XX DB 841 AGC 843
XX
XX .RESULT 2
XX AA236914
XX ID AA236914 standard; DNA; 1740 BP.
XX
XX AC AA236914;
XX
XX DT 13-MAR-2000 (first entry)
XX
XX DE DNA encoding an activator of G protein signalling (AGS) protein.
XX
XX KW Activator of G protein signalling; AGS; ras-related G protein;
XX GTP hydrolysis; G protein activity; pheromone response pathway;
XX G protein-coupled signal transduction; G-gamma selectivity;
XX cellular signal transduction; ss.
XX
XX KW Homo sapiens.
XX
XX OS
XX
XX FH key Location/Qualifiers
XX FT CDS 145..991
XX FT /*tag= a
XX FT /product= "activator of G protein signalling (AGS)
XX FT protein"
XX
XX Mo9958670-A1.
XX
XX PD 18-NOV-1999.
XX
XX PF 07-MAY-1999; 99WO-US10151.
XX
XX PR 08-MAY-1998; 98US-0084842.
XX 07-OCT-1998; 98US-0103355.
XX
XX PA (CADU-) CADUS PHARM CORP.

```

XX Cismowski M, Duzic E;
 XX MPI: 2000-072337/06.
 DR P-PSDB; AAY53924.
 XX
 PT A new activator of G protein signalling used to treat disorders
 PT characterized by an aberrant AGS protein activity -
 XX
 PS Disclosure; Page 146-148; 162pp; English.
 CC
 CC The present sequence encodes an activator of G protein signalling (AGS)
 CC protein. The cDNA sequence was isolated from a human liver cDNA
 CC library. The AGS protein exhibits homology to ras-related G proteins,
 CC and contains alterations in conserved amino acids consistent with a
 CC deficiency in GTP hydrolysis activity. AGS stimulates G protein
 CC activity, G protein-coupled signal transduction and the pheromone
 CC response pathway in a receptor-independent manner. The AGS protein
 CC also shows G-gamma selectivity, as measured by growth assays in
 CC yeast expressing various mammalian G-gamma constructs, and
 CC tissue-specific expression, as measured by Northern blot analysis.
 CC The AGS protein can be used to screen for compounds that modulate
 CC cellular signal transduction. The protein is used to treat disorders
 CC characterized by an aberrant AGS protein activity or AGS nucleic acid
 CC expression.
 CC
 SQ Sequence 1740 BP; 422 A; 546 C; 483 G; 289 T; 0 other;
 Alignment Scores:
 Pred. No.: 1,93e-157 Length: 1740
 Score: 1449.00 Matches: 281
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 21 Gaps: 0
 US-09-709-103-2 (1-281) x AA236914 (1-1740)
 QY 1 MetLysLeuAlaMetLysLysMetCysProSerAspSerGluLeuSerIlePro 20
 DB 146 ATGAAATCGCGCGGATGATCAAGAAAGTGTGCCGAGACACGAGCTGATATCCG 205
 QY 21 AAlaLysAncCysTyrArgMetValIleLeuGlySerSerLysValGlyThrAlaIle 40
 DB 206 GCCAAGAACACTGATCATGATGATCTCGCTCGCAAGGTGGGCAAGAGCGCCATC 265
 QY 41 ValSerArgPheLeuThrGlyArgPheGluAspAlaTyrThrProThrIleGluAspPhe 60
 DB 266 GTGTGCGGCTTCCACCGCGCGCTTCGAGAGACGCTACACGGCTACCATGAGAGACTTC 325
 QY 61 HisArgLysPheTyrSerIleArgGlyGluValTyrGlnLeuAspIleLeuAspThrSer 80
 DB 326 CACCGCAAGTTCTACTCATCCGCGGAGAGTGTACACAGTCGACATCTCGACAGCTCC 385
 QY 81 GlyAsnHisProPheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIle 100
 DB 386 GGCACACACCGCTTCCCGGCATGCGCGCTCTCCATCTCACAGAGACGTTTTCATC 445
 QY 101 LeuValPheSerLeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGlnGln 120
 DB 446 CTGCTGTTCAGTCTGACCAACCGGACTCTTCGAGAGAGTGCACGGCTCAGGACAG 505
 QY 121 IleLeuAspThrLysSerCysIleuLysAsnLysThrLysGluAsnValAspValProLeu 140
 DB 506 ATCTCGACACCAAGTCTTGCTCAAGAACAAACCAAGGAAGCGTGGAGCTGCCCTG 565
 QY 141 ValIleCysGlyAsnLysGlyAspArgAspPheTyrArgGluValAspGlnArgGluIle 160
 DB 566 GTCAATCTGCGGCAACAGGAGTACCGGACTTACCGCGAGGAGCAACGCGGAGATC 625
 QY 161 GluGlnLeuValGlyAspAspProGlnArgCysAlaTyrPheGluIleSerAlaLysLys 180
 DB 626 GAGCAGCTGGTGGGCGAGACCCCGCAGCGCTGCGCTACTTCGAGATCTCGCCCAAGAG 685

QY 181 AsnSerSerLeuAspGlnMetPheArgAlaLeuPheAlaMetAlaLysLeuProSerGlu 200
 DB 686 AACACAGAGCTTGACACAGATGTTCGCGCGCTCTCGCATATGCCAAGCTCCAGCGAG 745
 QY 201 MetSerProAspLeuHisArgLysValSerValGlnTyrCysAspValLeuHisLysLys 220
 DB 746 ATGACCCCAAGACCTTCACCGCAAGGTCTCGGTCAGTACTCGAGCTCTCACAAAGAG 805
 QY 221 AlaLeuArgAsnLysLysLeuLeuArgAlaGlySerGlyGlyGlyValAspProGly 240
 DB 806 GCGCTGCGGACAAAGAGTGTCTGCGGCGCGGACGCGCGCGCGCGGACCGGCG 865
 QY 241 AspAlaPheGlyIleValAlaProPheAlaArgArgProSerValHisSerAspLeuMet 260
 DB 866 GACGCTTGGCATCTGTGGACACCTTCGCGGCGCGGCGCGGACGATCACAGGACTCATG 925
 QY 261 TyrIleArgGluLysAlaSerLysArgGlnAlaLysAspLysGluArgCysValIle 280
 DB 926 TACATCCGCGAAGAGCGACGCGCGCAAGCCAGGCAAGAGCAAGAGCGCTGCTATC 985
 QY 281 Ser 281
 DB 986 AGC 988
 RESULT 3
 AA236893
 ID AA236893 standard; cDNA; 1801 BP.
 XX
 AC AA236893;
 XX
 DF 13-MAR-2000 (first entry)
 XX
 DE cDNA encoding an activator of G protein signalling (AGS) protein.
 XX
 KW Activator of G protein signalling; AGS; ras-related G protein;
 KW GTP hydrolysis; G protein activity; pheromone response pathway;
 KW G protein-coupled signal transduction; G-gamma selectivity;
 KW cellular signal transduction; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FH 5'UTR 7..153
 FT /*tag= a
 FT CDS 154..999
 FT /*tag= b
 FT /product= "activator of G protein signalling (AGS)
 FT 3'UTR 1000..1801
 FT /*tag= c
 PN WO9958670-A1.
 PD 18-NOV-1999.
 XX
 PF 07-MAY-1999; 99WO-US10151.
 XX
 PR 08-MAY-1998; 98US-0084842.
 PR 07-OCT-1998; 98US-0103355.
 XX
 PA (CADU-) CADUS PHARM CORP.
 XX
 PI Cismowski M, Duzic E;
 XX MPI: 2000-072337/06.
 DR P-PSDB; AAY53924.
 XX
 PT A new activator of G protein signalling used to treat disorders
 PT characterized by an aberrant AGS protein activity -
 PS Claim 3; Page 133-135; 162pp; English.
 XX

CC The present sequence encodes an activator of G protein signalling (AGS)
 CC protein. The cDNA sequence was isolated from a human liver cDNA
 CC library. The AGS protein exhibits homology to ras-related G proteins,
 CC and contains alterations in conserved amino acids consistent with a
 CC deficiency in GTP hydrolysis activity. AGS stimulates G protein
 CC response pathway in a receptor-independent manner. The AGS protein
 CC also shows G-gamma selectivity, as measured by growth assays in
 CC yeast expressing various mammalian G-gamma constructs, and
 CC tissue-specific expression, as measured by Northern blot analysis.
 CC The AGS protein can be used to screen for compounds that modulate
 CC cellular signal transduction. The protein is used to treat disorders
 CC characterized by an aberrant AGS protein activity or AGS nucleic acid
 CC expression.

XX Sequence 1801 BP; 437 A; 561 C; 500 G; 303 T; 0 other;

Alignment Scores:

Pred. No.:	2,02e-157	Length:	1801
Score:	1449.00	Matches:	281
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	21	Gaps:	0

US-09-709-103-2 (1-281) x AA236893 (1-1801)

```

QY 1 MetLysLeuAlaIaMeTLeLysLysMetCysProSerAspSerGluLeuSerIlePro 20
Db 154 ATGAACCTGCGCGATGATCAGAGATGTGCCGAGCGACTGAGATATCCG 213
QY 21 AlAlaLysCysTyrArgMetValIleLysSerSerLysValGlyLysThrAlaIle 40
Db 214 GCCAAGAACTCTATGCGATGTCATCTCGCTGTCGCAAGGTCGCAAGCGGCATC 273
QY 41 ValSerArgPheLeuThrGlyArgPheGluAspAlaTyrThrProTrIleGluAspHe 60
Db 274 GTGTGCGGCTTCTCACCAGCGCGCTTGAGAGACCTTACACGCTTACATGAGACTTC 333
QY 61 HisArgLysPheTyrSerIleArgGlyLysValIleGlnLeuAspIleLeuAspThrSer 80
Db 334 CACCGCAAGTCTACTCCATCCGCGGAGGTCTACAGCTCGACATCTCGACAGCTCC 393
QY 81 GlyAsnHisProPheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIle 100
Db 394 GGCACACCCGTTCCCGGCGATCGCGCTCTCCATCTCCACAGGAGAGAGCTTTTCATC 453
QY 101 LeuValPheSerLeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGlnGln 120
Db 454 CTGGTGTCAAGTGTGACACACCGGACTCTCTGAGAGAGGTGCAGCGGCTCAGGACGAG 513
QY 121 IleLeuAspThrLysSerCysLeuLysAsnLysThrLysGluAsnValAspValProLeu 140
Db 514 ATCTTCACACCAAGTCTTCTCCACAGCAAAACCAAGGAGAAACGTCGAGCTGCCCTG 573
QY 141 ValIleCysGlyAsnLysGlyAspArgAspPheTyrArgGluValAspGlnValIle 160
Db 574 GTCTATGCGGCAACAGGAGGTGACCGGACTTCTACCGGAGGTCGACCGCGCGAGATC 633
QY 161 GluGlnLeuValGlyAspAspProGlnArgCysAlaTyrPheGluIleSerAlaLysLe 180
Db 634 GAGAGAGTGTGGGCGAGACCCCGACGCTGCGCTACTTCAGATCTGCGGCAAGAG 653
QY 181 AsnSerSerLeuAspGlnMetPheArgAlaLeuPheAlaMetAlaLysLeuProSerGln 200
Db 694 AACAGCGCTGAGACAGATGTTCCGCGCTCTTCGCCATGCGCAAGCTGCCAGCGAG 753
QY 201 MetSerProAspLeuHisArgLysValSerValGlnTyrCysAspValLeuHisLysLys 220
Db 754 ATGAGGCCAAGCTGCGACCGCAAGGTCTCGGTCAAGTCTGCGCAAGCTGCGCAAGAG 813
QY 221 AlaLeuArgAsnLysLysLeuLeuArgAlaGlySerGlyGlyGlyGlyGlyGlyGlyGly 240

```

```

Db 814 GCGCTGCGAACAAGAGTGTGCGGCGCGGACGCGGCGCGGCGGCGACCGGCG 873
QY 241 AspAlaPheGlyIleValAlaProPheAlaIleArgProSerValHisSerAspLeuMet 260
Db 874 GACGCTTTGGCATCTGCGACCTTTCGCGCGGCGCGGCGGCGGCGGCGGCGGCGG 933
QY 261 TyrIleArgLysLysAlaSerAlaGlySerGlnAlaLysAspLysGluArgCysValIle 280
Db 934 TACATCGCGCAGAGAGCGGACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 993
QY 281 Ser 281
Db 994 AGC 996

```

RESULT 4

AA23022 standard; cDNA; 1841 BP.
 AC AA23022;
 DT 17-JAN-2000 (first entry)
 DE Human kd312 polypeptide encoding cDNA.

KW kd312 polypeptide; cancer; immunodeficiency disorder; AIDS; stroke;
 KW heart attack; head trauma; neurodegenerative disease; human;
 KW Parkinson's disease; Alzheimer's disease; ss.

OS Homo sapiens.

PN WO950288-A2.

PD 07-OCT-1999.

PF 30-MAR-1999; 99WO-US06693.

PR 31-MAR-1998; 98US-0053374.

PA (AMGE-) AMGEN INC.

PI Yen K;

XX WPI: 1999-601322/51.

DR P-PSDB: MAY42693.

PT kd312 polypeptides useful for treating diseases and disorders

XX associated with alterations in cell proliferation and cell death

PS Claim 1; Fig 8; 85pp; English.

CC The invention provides nucleic acid molecules encoding human and rat
 CC kd312 polypeptides. The kd312 polypeptides can be expressed by standard
 CC recombinant methodology. The kd312 sequences, and the antibodies against
 CC the proteins may be used to treat or diagnose the presence or progression
 CC of various diseases, e.g. cancer, immunodeficiency disorders (e.g. AIDS),
 CC stroke, heart attack, head trauma, and neurodegenerative diseases (e.g.
 CC Parkinson's disease and Alzheimer's disease). The present sequence
 CC represents the human kd312 cDNA sequence.

SQ Sequence 1841 BP; 398 A; 605 C; 530 G; 308 T; 0 other;

Alignment Scores:

Pred. No.:	2,09e-157	Length:	1841
Score:	1449.00	Matches:	281
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	20	Gaps:	0

US-09-709-103-2 (1-281) x AA23022 (1-1841)

```

QY 1 MetLysLeuAlaIaMeTLeLysLysMetCysProSerAspSerGluLeuSerIlePro 20

```

Dd	225	ATGAACCGGCGCGATGATCAAGAAATGATGCGCCGAGCATCTCGAGAGTGAATATCCG	31		
Oy	21	AlAlaysaScySYrArqMetValIlleuGlySerSerIysValGlyIsthrAlaIle	40		
Dd	315	GCCAAAGAACTCTATCCATGTCATCTCTCGGCTCGCCCAAGGTGGGCAAGAGCGGCATC	374		
Oy	41	ValSerArgPheLeuThrGlyArgPheGluAspAlaIsthrProThrIleGluAspPhe	60		
Dd	375	GTGTGGCGCTTCCTCACCGGCGCTTCGAGAGCGCTTACAGCGCTTACCATCGAGGACTTC	434		
Oy	61	HisArgIysPheTYrSerIleArgGlyGluValTYrGluLeuAspIleLeuAspThrSer	80		
Dd	435	CACCGCAAGTCTTACTCATDCGGGGGAGAGTCTACACAGCTGCAGCATCTCCGACAGTCC	494		
Oy	81	GlyAsnHisProPheProAlaMetArgIleuSerIleLeuThrGlyAspValPheIle	100		
Dd	495	GGCAACCAACCCGTTCCCGGCATCGCGCGCTCTCCATCTCACAGAGACGCTTTTCATC	554		
Oy	101	LeuValPheSerIleuAspAsnArgAspSerPheGluIuValGlnArgIleuArgGlnGlu	120		
Dd	555	CTGGGTTCACATCTCGACCAACCGGAGCTCTTCCAGAGAGGTGACGCGGCTCAAGCAGCAG	614		
Oy	121	IleLeuAspThrIlySerCysIleuIlyAsnIlysthrIysGlyAsnValAspValProIeu	140		
Dd	615	ATCCTCGACACACCAAGTCTGCTCAAGAACAAACCAAGAGAACTGAGACGTGCCCTG	674		
Oy	141	ValIleCysGlyAsnIlySGlyAspArgAspPheTYrArgGluValAspGlnArgGluIle	160		
Dd	675	GTCATCTCGCGCAACAAGGGTGACCGCGACCTTACCGGACGAGGNGACCGCGAGATC	734		
Oy	161	GluGlnIleuValGlyAspAspProGlnArgCysValArgIsthrGluIleSerAlaIlyIys	180		
Dd	735	GAGCAGCTGTGGGCGAGACCCCGCAGCGCTGACTTCGAGATCTGGGCCAAGAAC	794		
Oy	181	AsnSerSerIleuAspGlnMetPheArgAlaIleuPheAlaMetAlaIlyIleProSerGlu	200		
Dd	795	AACAGCAGCTCGGACAGCATGTTCCGGCGCTCTTCCGCATGGCCAAAGCTGCCACGAG	854		
Oy	201	MetSerProAspLeuHisArgIlyValSerValGlnTYrCysAspValIleuHisIlyIys	220		
Dd	855	ATGAGCCCAAGACCTGACACCGCAAGGTCTCGGTGCAAGTACGCGACGTGTGCACAAGAC	914		
Oy	221	AlaIleuArgAsnIlyIysLeuIleuArgAlaIlySerGlyValGlyGlyGlyAspProGly	240		
Dd	915	GCGCTCGCGAACAAGAACTCTCTCGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	974		
Oy	241	AspAlaPheGlyIleValAlaProPheAlaIsthrArgIsthrProSerValHisSerAspLeuMet	260		
Dd	975	GAGCGCTTTGGATCTGTGACCCCTTGGCGGGCGCGCGCGCGCGCGCGCGCGCGCGCG	1030		
Oy	261	TYrIleArgGlyIlyAspAlaSerAlaGlySerGlnAlaIlyAspIlySGlyIsthrCysValIle	280		
Dd	1035	TACATCTCGCGAAGAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	1090		
Oy	281	Ser 281			
Dd	1095	AGC 1097			
RESULT 5					
ID	AAA49177	standard; cDNA: 1776 BP.			
XX	AAA49177;				
XX	03-NOV-2000	(first entry)			
DE	cDNA encoding human GTPase associated protein-7.				
KW	Guanine nucleotide binding protein; GTP-binding protein; G-protein;				
KW	GTPase; GTPase associated protein; GTPAP; cell proliferation;				
KW	autoimmune; inflammatory; immune system disorder; cancer; AIDS;				
KW	acquired immune deficiency syndrome; asthma; atherosclerosis;				
KW	arthritis; systemic lupus erythematosus; psoriasis; human; ss.				

QY 81 GlyAsnHisProPheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIle 100
 DB 420 GGCAACACACCCGTCGCCGACATGCGGTGCTCTCATCTCCACAGAGAGGTTTTCATC 479
 QY 101 LeuValPheSerLeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGln 120
 DB 480 CTGGGTTCAGTCTGGACACCGGACCTCTTCGAGGAGGTGACAGCGGCTCAGCGAGAG 539
 QY 121 IleLeuAspThrLysSerCysLeuLysAsnLysThrLysGluAsnValAspValProLeu 140
 DB 540 ATCCCTGCAGACCAACTCTTCCCTCAAGAACAAACCAAGAGAGAGCTGACCTGCCCCCTG 599
 QY 141 ValIleCysGlyAsnLysGlyAspArgAspPheThrArgGluValAspGlnArgGluIle 160
 DB 600 GTCATCTCGCGGACACAGGAGTGACCGGACTTCTACCGCGAGGTGAGACACGCGAGATC 659
 QY 161 GluGlnLeuValGlyAspAspProGlnArgCysAlaThrPheGluIleSerAlaLysLys 180
 DB 660 GAGCAGCTGTGTGGCGACGACCCCGAGGCTGGCGCTTCTGAGATCTCGGCCAAGAGAG 719
 QY 181 AsnSerSerLeuAspGlnMetPheArgAlaLeuPheAlaMetAlaLysLeuProSerGlu 200
 DB 720 AACAGACACCTGCAGACCAATGTTCGCCGCTCTTCCGACATGGCCAGACTGCCAGCGAG 779
 QY 201 MetSerProAspLeuHisArgLysValSerValGlnThrCysAspValLeuHisLysLys 220
 DB 780 ATGAGCCCGAGCTGCACCGCAGAGGTCTCGGTGCAAGTACTCGAGCTGCTCACAAGAG 839
 QY 221 AlaLeuArgAsnLysLysLeuLeuArgAlaGlySerGluGluGluGluAspProGlu 240
 DB 840 GCGCTGCGGAAACAGAAAGCTCTGCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 899
 QY 241 AspAlaPheGlyIleValAlaProPheAlaArgArgProSerValHisSerAspLeuMet 260
 DB 900 GACGCTTGGCATCGTGGACACCTTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 959
 QY 261 TyrIleArgGluLysAlaSerAlaGlySerGlnAlaLysAspLysGluArgCysValIle 280
 DB 960 TACATCCCGGAGAAAGGCGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1019
 QY 281 Ser 281
 DB 1020 AGC 1022
 RESULT 6
 AA23024
 ID AA23024 standard; cDNA; 1689 BP.
 AC AA23024;
 XX
 DT 17-JAN-2000 (first entry)
 XX
 DE Rat kd312 polypeptide encoding cDNA.
 XX
 KW kd312 polypeptide; cancer; immunodeficiency disorder; AIDS; stroke;
 KW heart attack; head trauma; neurodegenerative disease; rat;
 KW Parkinson's disease; Alzheimer's disease; ss.
 XX
 OS Rattus sp.
 XX
 PN MO9950288-A2.
 XX
 PD 07-OCT-1999.
 XX
 PF 30-MAR-1999; 99MO-US06993.
 XX
 PR 31-MAR-1998; 98US-0053374.
 XX
 PA (AMGE-) AMGEN INC.
 XX
 PI Yen K;
 XX

DR WP1; 1999-601322/51.
 DR P-RSDB; AA42694.
 XX
 PT kd312 polypeptides useful for treating diseases and disorders
 PT associated with alterations in cell proliferation and cell death
 XX
 PS Claim 2; Fig 9; 85pp; English.
 XX
 CC The invention provides nucleic acid molecules encoding human and rat
 CC kd312 polypeptides. The kd312 polypeptides can be expressed by standard
 CC recombinant methodology. The kd312 sequences, and the antibodies against
 CC of various diseases, e.g. cancer, immunodeficiency disorders (e.g. AIDS),
 CC stroke, heart attack, head trauma, and neurodegenerative diseases (e.g.
 CC Parkinson's disease and Alzheimer's disease). The present sequence
 CC represents the rat kd312 cDNA sequence.
 XX
 SQ Sequence 1689 BP; 417 A; 462 C; 434 G; 376 T; 0 other;
 Alignment Scores:
 Pred. No.: 9, 87e-152 Length: 1689
 Score: 1399.50 Matches: 273
 Percent Similarity: 98.228 Conservative: 3
 Best Local Similarity: 97.158 Mismatches: 4
 Query Match: 96.58% Indels: 1
 DB: 20 Gaps: 1
 US-09-709-103-2 (1-281) x AA23024 (1-1689)
 QY 1 MetLysLeuAlaAlaMetIleLysLysMetCysProSerAspSerGluLeuSerIlePro 20
 DB 132 ATGAACTGGCCCCGATGATCAAGAGATGTGCCCAAGGACCTGAACTGATTCGCCG 191
 QY 21 AlaLysAspCysThrArgMetValIleLeuGlySerSerLysValGlyLysThrAlaIle 40
 DB 192 GCCAAGACTGCTACAGAGATGTCATCTCGGCTCATCCAAAGTGCGCAAGAGGCCATC 251
 QY 41 ValSerArgPheLeuThrGlyArgPheGluAspAlaThrThrProThrIleGluAspHe 60
 DB 252 GTGTCGCGCTCTCTCAAGGCGGCGCTCGAGGAGCGCTTACACCCCTTACAACTTC 311
 QY 61 HisArgLysPheThrSerIleArgGlyGluValTyrGlnLeuAspIleLeuAspThrSer 80
 DB 312 CACGAAAGTTTACTTCGATCCGCGGCGGCAAGCTTACACCTTGACATCTGACACACT 371
 QY 81 GlyAsnHisProPheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIle 100
 DB 372 GGCAATCATCCGTTTCCGCCCATGCGGCGCTCTCTATCTTCAAGAGAGCGTTTTCAT 431
 QY 101 LeuValPheSerLeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGln 120
 DB 432 CTGGTGTTCAGCTTAGACAAACGCGACTCTTCGAGGAGGTGCAAAAGGCTCAACAGCAG 491
 QY 121 IleLeuAspThrLysSerCysLeuLysAsnLysThrLysGluAsnValAspValProLeu 140
 DB 492 ATCCCTGCAGACCAACTCTTCCCTCAAGAACAAACCAAGAGAGTGTGACCTGCCGCTG 551
 QY 141 ValIleCysGlyAsnLysGlyAspArgAspPheThrArgGluValAspGlnArgGluIle 160
 DB 552 GTCATCTCGCGGACACAGGAGTGACCGGACTTCTACCGGAGGTGAGAGCGGAGATT 611
 QY 161 GluGlnLeuValGlyAspAspProGlnArgCysAlaThrPheGluIleSerAlaLysLys 180
 DB 612 GAGCAGCTGTGTGGCGACGATGACCTCGGCTGTGCTACTTGAATCTCGGCCAAGAGAG 671
 QY 181 AsnSerSerLeuAspGlnMetPheArgAlaLeuPheAlaMetAlaLysLeuProSerGlu 200
 DB 672 AATGAGAGCTTGAGCAGATGTTCCGCGCTTTCGCAATGCGCAACCTGCTGCGAG 731
 QY 201 MetSerProAspLeuHisArgLysValSerValGlnThrCysAspValLeuHisLysLys 220
 DB 732 ATGAGCCCTGACTTGACCGCAAGGTGTCTGTGCAAGTACTGTGACCTGTGCAACAAAG 791

QY 221 AlaLeuArgAsnLysLysLeuLeuArgAlaGlySerGlyGlyGlyGlyAspProGly 240
 Db 792 GCTCTGTGGAGCAAGAGCTTCTGCGTGGGCGAGC---GGAGGTGGGGCGCCACCGA 848
 QY 241 AspAlaPheGlyIleValAlaProPheAlaArgProSerValHisSerAspLeuMet 260
 Db 849 GATGCGCTTGGCATCTTGGCGCCCTTGTCTGCAACCTAAGCTGCAATGCACTCATG 908
 QY 261 TyrIleArgGlyLysLysAlaSerAlaGlySerGlnAlaLysAspLysGlnArgCysValIle 280
 Db 909 TACATTCGTGAGAAACCAAGCTGTCAGCAGCCAGCCTAAGAGACAAGAGCGCTGTGCATC 968
 QY 281 Ser 281
 Db 969 AGT 971

RESULT 7
 AA223023
 ID AA223023 standard; DNA: 3986 BP.

AC AA223023;

DT 17-JAN-2000 (first entry)

XX Human kd312 genomic DNA sequence.

XX kd312 polypeptide; cancer; immunodeficiency disorder; AIDS; stroke;

KW heart attack; head trauma; neurodegenerative disease; human;

KW Parkinson's disease; Alzheimer's disease; ss.

OS Homo sapiens.

PN MO950288-A2.

PD 07-OCT-1999.

PF 30-MAR-1999; 99WO-US06993.

PR 31-MAR-1998; 98US-0053374.

XX (AMGE-) AMGEN INC.

PI Yen K.

XX WPI: 1999-601322/51.

DR P-PSDB: AAY42693.

XX

PT kd312 polypeptides useful for treating diseases and disorders

PT associated with alterations in cell proliferation and cell death

PS Claim 1; Fig 5; 85pp; English.

XX

XX The invention provides nucleic acid molecules encoding human and rat

CC kd312 polypeptides. The kd312 polypeptides can be expressed by standard

CC recombinant methodology. The kd312 sequences, and the antibodies against

CC the proteins may be used to treat or diagnose the presence or progression

CC of various diseases, e.g. cancer, immunodeficiency disorders (e.g. AIDS),

CC stroke, heart attack, head trauma, and neurodegenerative diseases (e.g.

CC Parkinson's disease and Alzheimer's disease). The present sequence

CC represents the human kd312 genomic DNA sequence.

XX

Sequence 3986 BP; 856 A; 1226 C; 1172 G; 732 T; 0 other.

Alignment Scores:

Pred. No.: 9.36e-150

Score: 1387.00

Percent Similarity: 79.83%

Best Local Similarity: 79.83%

Query Match: 95.72%

DB: 20

Length: 3986

Matches: 281

Conservative: 0

Mismatches: 71

Gaps: 1

US-09-709-103-2 (1-281) x AA223023 (1-3986)

QY 1 MetLysLeuAlaIleMetIleLysMetCysProSerAspSerGlnLeuSerIlePro 20
 Db 774 ATGAACATGGCCGCGGATGATCAAGAGATGTCGCCAGCATCTGGACCTAGATATCCG 833
 QY 21 AlaLysAsnCysTyrArgMetValIleLeuGlySerSerLysValGlyLysThrAlaIle 40
 Db 834 GCCAAGACGTGTATCGCATGTGTCATCTCGCTGTGTCCAAAGTGGGCAACGCGCATC 893
 QY 41 ValSerArgPheLeuThrGlyArgPheGlnAspAlaTyrThrProThrIleGlnAspPhe 60
 Db 894 GTGTGCGCTTCTCAGCGCGCTTCAGAGACGCTACAGCCCTACCATCAGAGACTTC 953
 QY 61 HisArgLysPheTyrSerIleArgGlyIleValTyrGlnLeuAspIleLeuAspThrSer 80
 Db 954 CACCGCAAGTTCTACTCCATCCGCGGAGAGTCTACAGCTGACATCTCGACACTTC 1013
 QY 81 GlyAsnHisProPheProAlaMetArgArgLeuSerIleLeuThr----- 95
 Db 1014 GGCAACCAACCGCTTCCCGCCATGCGGCGCTCTCATCTCTCAGAGTGAGCGGGGCGC 1073
 QY 95 ----- 95
 Db 1074 GGGCAGGTGCGGGAGGAGGAGGCGGGGAACCTCGGCCAGGGCGCCCGCAGCCGCT 1133
 QY 95 ----- 95
 Db 1134 CCGGCTGCTGCGCGCCGAGTACTGCGCTTCCGCTTACAGAGCTTACGCGCCGCGCGC 1193
 QY 95 ----- 95
 Db 1194 GCCTCAAAAGTCAGCCGAGTCTTCCCTGGGCGCCACCTTCACCTTCTCTCTCT 1253
 QY 96 -----GlyAspValPheIleLeuValPheSerLeuAspAsnArgAspSe 110
 Db 1254 CTTCTGCCCCCTCTAGAGAGAGCTTTTCACTCTGCTTCACTGAGCAACCGGAGCTC 1313
 QY 110 rPheGlnGluValGlnArgLeuArgGlnGlnIleLeuAspThrLysSerCysLeuLysAs 130
 Db 1314 CTTGAGAGAGTGCAGCGCTCAGCGAGAGATCTCGACACCAAGTCTGCTTCAAGAA 1373
 QY 130 nLysThrLysGluAsnValAspValProLeuValIleCysGlyAsnLysGlyAspArgAs 150
 Db 1374 CAACCAAGAGAGAGTGTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1433
 QY 150 PheTyrArgGluValAspGlnArgGluIleGlnIleLeuValGlyAspAspProGlnArg 170
 Db 1434 CTTCTACCGCGAGGTGAGCGAGCGAGATCGACAGCTGTGGGGGAGACCCCGAGCG 1493
 QY 170 GlysAlaTyrPheGlnIleSerAlaLysLysAsnSerSerLeuAspGlnMetPheArgAl 190
 Db 1494 CTGGCGCTTCTTGGAGATCTCGGCGCAAGAGACAGACAGCTGTGAGCGAGATGTCGCGC 1553
 QY 190 AleuPheAlaMetAlaLysLeuProSerGlnMetSerProAspLeuHisArgLysValSe 210
 Db 1554 GCTCTGCGCATGCGCAAGTGTGCGCGAGATGAGACCCAGACCTTCCAGCAAGGTCTC 1613
 QY 210 rValGlnTyrCysAspValLeuHisLysLysAlaLeuAsnLysLysLeuLeuArgAl 230
 Db 1614 GGTGAGTACTGCGAGCTGTGCGACAAAGAGGCGTGGGCAACAAACCTCTGCGGGC 1673
 QY 230 agLysSerGlyGlyGlyGlyGlyAspProGlyAspAlaPheGlyIleValAlaProPheAl 250
 Db 1674 CGGCAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1733
 QY 250 ArgArgProSerValHisSerAspLeuMetTyrIleArgGlyLysAlaSerAlaGlySe 270
 Db 1734 GCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1793
 QY 270 rGlnAlaLysAspLysGlnArgCysValIleSer 281
 Db 1794 CCAAGCCCAAGAGACAGAGCGCTGCGTCAATCAGC 1827

RESULT 8

AA23025
 ID AA23025 standard; DNA; 3079 BP.
 AC AA23025;
 DT 17-JAN-2000 (first entry)
 XX
 DE Rat kd312 genomic DNA sequence.
 XX
 KW kd312 polypeptide; cancer; immunodeficiency disorder; AIDS; stroke;
 KW heart attack; head trauma; neurodegenerative disease; rat;
 KW Parkinson's disease; Alzheimer's disease; ss.
 XX
 OS Rattus sp.
 PN MO9950288-A2.
 PD 07-OCT-1999.
 PF 30-MAR-1999; 99WO-US06993.
 PR 31-MAR-1998; 98US-0053374.
 PA (AMGE-) AMGEN INC.
 PY Yen K;
 DR WPI: 1999-601322/51.
 DR P-PSDB: AAY42694.
 XX
 PT kd312 polypeptides useful for treating diseases and disorders
 XX associated with alterations in cell proliferation and cell death
 XX
 PS Claim 2: Fig 7: 85pp; English.
 XX
 CC The invention provides nucleic acid molecules encoding human and rat
 CC kd312 polypeptides. The kd312 polypeptides can be expressed by standard
 CC recombinant methodology. The kd312 sequences, and the antibodies against
 CC the proteins may be used to treat or diagnose the presence or progression
 CC of various diseases, e.g. cancer, immunodeficiency disorders (e.g. AIDS),
 CC stroke, heart attack, head trauma, and neurodegenerative diseases (e.g.
 CC Parkinson's disease and Alzheimer's disease). The present sequence
 CC represents the rat kd312 genomic DNA sequence.
 XX
 SQ Sequence 3079 BP: 708 A; 843 C; 829 G; 699 T; 0 other;
 Alignment Scores:
 Pred. No.: 7.99e-145 Length: 3079
 Score: 1343.00 Matches: 272
 Percent Similarity: 84.10% Conservative: 3
 Best Local Similarity: 83.18% Mismatches: 5
 Query Match: 92.68% Indels: 48
 Gaps: 2
 US-09-709-103-2 (1-281) x AA23025 (1-3079)
 QY 1 MetLysLeuAlaMetLLeuLysLysMetCysProSerAspSerGluLeuSerIlePro 20
 DB 303 ATGAACCTGGCGCGATGATCAGAGATGCGCCAAAGCACTGTAACAGATATCCG 362
 QY 21 AlAlysAnCyTYrArGmetValIleuGlySerSerLysValGlyLThrAlaIle 40
 DB 363 GCCAAGAACTGCTACAGATGCTATCTCGGCTCATCCAAAGTGGGCAAGAGCGCCATC 422
 QY 41 ValSerArGheLeuThrGlyArGpHeGluAspAlaTyThrProThrIleGluAspPhe 60
 DB 423 GTGTCCGCTTCCCTCACGGCGCGCTTCGAGAGACCTTACACCCCTACCATTTGAAGCTTC 482
 QY 61 HisArgLysPheTySerIleArgGlyGluValTyrcLLeuAspIleLeuAspThrSer 80
 DB 483 CACCGAAAGTTTACTGATCCGGGCGAAGCTACACAGTTGACATACAGACATCT 542
 QY 81 GlyAsnHisProPheProAlaMetArgArgLeuSerIle----- 93

DB 543 GGCATATCATCGCTTCCCGGCATCGCGCTCTCTATCTCCTACAGAGTGAAGTGGGACC 602
 QY 93 ----- 93
 DB 603 GACAGGACCGTGGGAGGAATCTCGGGGAGCGGATGGGGCGTGTGTGTGG 662
 QY 94 -----Leu 94
 DB 663 GCTGTGCTGTGCTGCTCCGTGCTGGAGTGCCTCACCCTTTCACCTGTTCCCTTG 722
 QY 95 ThrGlyAspValPheIleLeuValPheSerLeuAspAsnArgAspSerPheGluVal 114
 DB 723 TA-CGAGACGTTTTCATTTGTTGTTGAGTGTGACACCGGACCTCTTCGAGGAGT 781
 QY 115 GluArgLeuArgGlnGlnIleLeuAspThrLysSerCysLeuLysAsnLysThrLysGlu 134
 DB 782 CAAAGGCTCAACAGCAGATCTAGACACCAAGTCTCTCAAGAACAAACCAAGAG 841
 QY 135 AsnValAspValProLeuValIleCysGlyAsnLysGlyAspArgAspPheTyArgGlu 154
 DB 842 AATGTGACGTGCGCGCTGTCATTTGGGGTACAAAGGGAGCGGACTTCTACCGCGAA 901
 QY 155 ValAspGlnArgGluIleGluGlnLeuValGlyAspAspProGlnArgCysAlaTyPhe 174
 DB 902 GTGAGCAGCAGCGGAGATTGAGCAGCTGTGGCGGAGATGACCGCTTGCTGCTACTTC 961
 QY 175 GluIleSerLysLysAsnSerSerLeuAspGlnMetPheArgAlaLeuPheAlaMet 194
 DB 962 GAGATCTCGCGCAAGAAAGTATGACCGCTGACCAATGTCCTGCTCTTGGCCATG 1021
 QY 195 AlAlysLeuProSerGluMetSerProAspLeuHisArgLysValSerValGlnTyrcys 214
 DB 1022 GCCAAGCTGCTAGCGAGATGAGACCTTGTACCCGAAAGGTGTCTGTGACGATCTGT 1081
 QY 215 AspValLeuHisLysLysAlaLeuArgAsnLysLysLeuLeuArgAlaGlySerGlyGlu 234
 DB 1082 GACGTCTGTGCAAAAAGCTGTGAGGAACAAGAACTCTGCGTGGGCGAGC---GGA 1138
 QY 235 GlyGlyGlyAspProGlyAspAlaPheGlyTyrIleValIleProPheAlaArgArgProSer 254
 DB 1139 GTGGGGGCGACACCGAGAGTGGCTTGGCATCTTGGCGCCCTTTCCTGCGAGACCTTAC 1198
 QY 255 ValHisSerAspLeuMetTyrrIleArgGluLysAlaSerAlaGlySerGlnAlaLysAsp 274
 DB 1199 GTGCATAGCGACCTCATGTACATTCGTGAGAAACAGATGTACAGCAGCGCTAAGGAC 1258
 QY 275 LysGluArgCysValIleSer 281
 DB 1259 AAGGAGCGCTGTGTATCATG 1279
 RESULT 9
 AA236913
 ID AA236913 standard; DNA; 837 BP.
 AC AA236913;
 DT 13-MAR-2000 (first entry)
 XX
 DE DNA encoding a homologue of activator of G protein signalling AGS1.
 XX
 KW Activator of G protein signalling; AGS; ras-related G protein;
 KW GTP hydrolysis; G protein activity; pheromone response pathway;
 KW G protein-coupled signal transduction; G-gamma selectivity;
 KW cellular signal transduction; AGS1 homologue; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..837
 FT /*tag= a
 FT /*product= "AGS1 homologue"

CC tumour endothelial marker (TEM) protein selected from ABB90732, ABB90740,
 CC ABB90749, ABB90750 and ABB90769. The antibodies which bind to TEM
 CC proteins have cytostatic, immunostimulant and antiangiogenic activity.
 CC They are useful for inhibiting tumour growth, neoplasia, in
 CC subjects bearing a vascularised tumour, polycystic kidney disease,
 CC diabetic retinopathy, rheumatoid arthritis and psoriasis. Human, mouse
 CC and rat TEM genes and the encoded proteins (ABL92075-ABL92141 and
 CC ABB90721-ABB90789) are disclosed, as are marker oligonucleotide
 CC sequences, tumour endothelial markers (TEM) ABL91996-ABL92041 and
 CC ABL92143-ABL92191; normal endothelial markers (NEM) ABL92042-ABL92074;
 CC and pan-endothelial markers (PEM) ABL91903-ABL91995.

SQ Sequence 2832 BP; 602 A; 891 C; 727 G; 612 T; 0 other;

Alignment Scores:

Pred. No.:	8.3e-93	Length:	2832
Score:	893.00	Matches:	174
Percent Similarity:	78.14%	Conservative:	44
Best Local Similarity:	62.37%	Mismatches:	45
Query Match:	61.63%	Indels:	16
DB:	24	Gaps:	4

US-09-709-103-2 (1-281) x ABL92076 (1-2832)

```

OY 5 AlamelilelysluysmetcysproserasphergluLeuSerlleproalalyanycs 24
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 34 GCCATGATGAAGACTTTGTCACCGGAGAACCTGACGCTGCTGCGCCCAAAAATCA 93
OY 25 TTTATGethValleleuglySerSerlyslValglyThrAlaIleValSerarph 44
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 94 TACCGCATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 153
OY 45 LeuThrglyArphhegluaspalaTyrrhrprhrhrlegluasphehisarglyspe 64
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 154 CTCGAAGGGCGCTTTAGAGCACGATACACACCACCTGAGGACTTCCACCGTAAGGTA 213
OY 65 TyrrSerlleArglyslValTyrrgluLeuAspIleLeuAspThrSerGlyAsnHispro 84
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 214 TACACATCCGGCGGCGCATGTACACAGCTCGACATCTCGGACACACACCC 273
OY 85 PheProAlaMetArgrgluSerlleleuthrglyAspValPheIleLeuValPheSer 104
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 274 TTCCTCCGATGGCGAGGTGTCATCTTCACAGGAGGATGCTTCATCTGTTGTCAGC 333
OY 105 LeuAspAsnArgrAspSerPhegluValGlnArgrleuArgrGlnIleLeuAspThr 124
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 334 CTGGATTAACCGGAGTCTTCGATGAGGTCAAGCCCTTCAGAGCAGATCTCGAGGTC 393
OY 125 LysSerCysleuLysAsnLysThrLysGluAsnValAspValProleuValIleCysGly 144
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 394 AAGTCTGCTGTAAAGAACAGACAGGAGCGCGGCGGCGCCATGCTGATCTGAGGC 453
OY 145 AsnLysGlyAspArg---AspPheTyrrArgrgluValAspGlnArgrGluIleGluInleu 163
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 454 AACAAAGAACAGACAGGAGCTGCGGCGGAGGTGCGCCACACCGAGGCGGCTGCTG 513
OY 164 ValGlyAspAspProGlnArgrCysAlaTyrrPhegluIleSerAlaLysLysAsnSerSer 183
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 514 GTGTGGGGCGAC---GACAACTGCGCCCACTTCAGAGTGTGCGGCAAGAACACCAAC 570
OY 184 LeuAspGlnMetPheArgrAlaLeuPheAlaMetAlaLysLysLeuProSerGluMetSerPro 203
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 571 GTGAGCAGATGTTCTTACGCTGCTTCAGCATGGCCAAAGCTGCGCACAGATGAGCCCC 630
OY 204 AspLeuHisArgrLysValSerValGlnTyrrCysAspValLeuHisLysLys-----Ala 221
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 631 GCCCTGATCCAGATCTCGTCCAGTACGAGTACGCCCTTCACCCAGGCCCTTCTGCG 690
OY 222 LeuArgrAsnLysLysLeuLeuArgrAlaGlySerGlyCylGlyLysLysAspProGlyAsp 241
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 691 ATGGCGCCGCTCAAGAGATG-----GAC 714
OY 242 AlaPheGlyIleValAlaPhePheAlaArgrArgrProSerValHisSerAspLeuMetLys 261

```

Db 715 GCCTATGCAAGGTCTGCCCCCTTGCGCCGCGCCGCTCAACAGTGCATCAAGTAC 774

OY 262 lIeargrluysAlaSerAlaGlySerGlnAlaLysAspLysLysLysValIle 280

Db 775 ATCAAGGCCAAGGTCCTTGCGGAGGAGCCAGCCGCTGAGGAGCAAGTGCACCATC 831

RESULT 11

ABL92087

ID ABL92087 standard; cDNA; 2973 BP.

AC ABL92087;

DT 30-MAY-2002 (first entry)

DE Human Tumour Endothelial Marker polynucleotide seq ID NO 197.

KW Human; mouse; rat; TEM; tumour endothelial marker; NEM; PEM; cytostatic;
 KW normal endothelial marker; pan-endothelial marker; immunostimulant;
 KW antiangiogenic; tumour; neoplasia; diabetes; retinopathy; rheumatoid arthritis;
 KW polycystic kidney disease; diabetes; retinopathy; rheumatoid arthritis;
 KW psoriasis; gene; ss.

OS Homo sapiens.

PN WO200210217-A2.

PD 07-FEB-2002.

PF 01-ANG-2001; 2001WO-US24031.

PR 02-AUG-2000; 2000US-222599P.

PR 11-AUG-2000; 2000US-224360P.

PR 11-APR-2001; 2001US-282850P.

XX (UYTO) UNITV JOHNS HOPKINS.

XX St Croix B. Kinzler KW, Vogelstein B;

XX P-PSDB; ABB90733.

XX WPI; 2002-291856/33.

XX DR P-PSDB; ABB90733.

XX PT An isolated molecule comprising an antibody variable region which

PT specifically binds to an extracellular domain of a tumor endothelial

XX marker (TEM) protein, useful for inhibiting tumor growth

XX Claim 65; Page 152-153; 331pp; English.

XX The invention relates to an isolated molecule comprising an antibody

XX variable region which specifically binds to an extracellular domain of a

XX tumor endothelial marker (TEM) protein selected from ABB90732, ABB90740,

XX ABB90749, ABB90750 and ABB90769. The antibodies which bind to TEM

XX proteins have cytostatic, immunostimulant and antiangiogenic activity.

XX They are useful for inhibiting tumour growth, neoplasia, in

XX subjects bearing a vascularised tumour, polycystic kidney disease,

XX diabetic retinopathy, rheumatoid arthritis and psoriasis. Human, mouse

XX and rat TEM genes and the encoded proteins (ABL92075-ABL92141 and

XX ABB90721-ABB90789) are disclosed, as are marker oligonucleotide

XX sequences, tumour endothelial markers (TEM) ABL91996-ABL92041 and

XX ABL92143-ABL92191; normal endothelial markers (NEM) ABL92042-ABL92074;

XX and pan-endothelial markers (PEM) ABL91903-ABL91995.

SQ Sequence 2973 BP; 596 A; 967 C; 787 G; 623 T; 0 other;

Alignment Scores:

Pred. No.:	8.89e-93	Length:	2973
Score:	893.00	Matches:	174
Percent Similarity:	78.14%	Conservative:	44
Best Local Similarity:	62.37%	Mismatches:	45
Query Match:	61.63%	Indels:	16
DB:	24	Gaps:	4

US-09-709-103-2 (1-281) x ABL92087 (1-2973)

[illegible]

OS	Homo sapiens.		
XX	adrenal gland; bone marrow; lung; ovary; pancreas; prostate; spleen;		
KW	skin; testis; thymus.		
XX			
PN	WO200220754-A2.		
XX			
PD	14-MAR-2002.		
XX			
PE	29-AUG-2001; 2001WO-US27127.		
XX			
PR	05-SEP-2000; 2000US-229747P.		
PR	05-SEP-2000; 2000US-229748P.		
PR	05-SEP-2000; 2000US-229749P.		
PR	05-SEP-2000; 2000US-229750P.		
PR	05-SEP-2000; 2000US-229751P.		
PR	06-SEP-2000; 2000US-230583P.		
PR	06-SEP-2000; 2000US-230585P.		
PR	06-SEP-2000; 2000US-230586P.		
PR	06-SEP-2000; 2000US-230587P.		
PR	06-SEP-2000; 2000US-230588P.		
PR	06-SEP-2000; 2000US-230589P.		
PR	06-SEP-2000; 2000US-230590P.		
PR	06-SEP-2000; 2000US-230591P.		
PR	06-SEP-2000; 2000US-230592P.		
PR	06-SEP-2000; 2000US-230593P.		
PR	06-SEP-2000; 2000US-230594P.		
PR	06-SEP-2000; 2000US-230595P.		
PR	06-SEP-2000; 2000US-230596P.		
PR	06-SEP-2000; 2000US-230597P.		
PR	06-SEP-2000; 2000US-230598P.		
PR	06-SEP-2000; 2000US-230599P.		
PR	06-SEP-2000; 2000US-230600P.		
PR	06-SEP-2000; 2000US-230601P.		
PR	06-SEP-2000; 2000US-230602P.		
PR	06-SEP-2000; 2000US-230603P.		
PR	06-SEP-2000; 2000US-230604P.		
PR	06-SEP-2000; 2000US-230605P.		
PR	06-SEP-2000; 2000US-230606P.		
PR	06-SEP-2000; 2000US-230607P.		
PR	06-SEP-2000; 2000US-230608P.		
PR	06-SEP-2000; 2000US-230609P.		
PR	06-SEP-2000; 2000US-230610P.		
PR	06-SEP-2000; 2000US-230611P.		
PR	06-SEP-2000; 2000US-230612P.		
PR	06-SEP-2000; 2000US-230613P.		
PR	06-SEP-2000; 2000US-230614P.		
PR	06-SEP-2000; 2000US-230615P.		
PR	06-SEP-2000; 2000US-230616P.		
PR	06-SEP-2000; 2000US-230617P.		
PR	06-SEP-2000; 2000US-230618P.		
PR	06-SEP-2000; 2000US-230619P.		
PR	06-SEP-2000; 2000US-230620P.		
PR	06-SEP-2000; 2000US-230621P.		
PR	06-SEP-2000; 2000US-230622P.		
PR	06-SEP-2000; 2000US-230623P.		
PR	06-SEP-2000; 2000US-230624P.		
PR	06-SEP-2000; 2000US-230625P.		
PR	06-SEP-2000; 2000US-230626P.		
PR	06-SEP-2000; 2000US-230627P.		
PR	06-SEP-2000; 2000US-230628P.		
PR	06-SEP-2000; 2000US-230629P.		
PR	06-SEP-2000; 2000US-230630P.		
PR	06-SEP-2000; 2000US-230631P.		
PR	06-SEP-2000; 2000US-230632P.		
PR	06-SEP-2000; 2000US-230633P.		
PR	06-SEP-2000; 2000US-230634P.		
PR	06-SEP-2000; 2000US-230635P.		
PR	06-SEP-2000; 2000US-230636P.		
PR	06-SEP-2000; 2000US-230637P.		
PR	06-SEP-2000; 2000US-230638P.		
PR	06-SEP-2000; 2000US-230639P.		
PR	06-SEP-2000; 2000US-230640P.		
PR	06-SEP-2000; 2000US-230641P.		
PR	06-SEP-2000; 2000US-230642P.		
PR	06-SEP-2000; 2000US-230643P.		
PR	06-SEP-2000; 2000US-230644P.		
PR	06-SEP-2000; 2000US-230645P.		
PR	06-SEP-2000; 2000US-230646P.		
PR	06-SEP-2000; 2000US-230647P.		
PR	06-SEP-2000; 2000US-230648P.		
PR	06-SEP-2000; 2000US-230649P.		
PR	06-SEP-2000; 2000US-230650P.		
PR	06-SEP-2000; 2000US-230651P.		
PR	06-SEP-2000; 2000US-230652P.		
PR	06-SEP-2000; 2000US-230653P.		
PR	06-SEP-2000; 2000US-230654P.		
PR	06-SEP-2000; 2000US-230655P.		
PR	06-SEP-2000; 2000US-230656P.		
PR	06-SEP-2000; 2000US-230657P.		
PR	06-SEP-2000; 2000US-230658P.		
PR	06-SEP-2000; 2000US-230659P.</		

QY 195 AlAlYsLeuPProSerGluMetSerProAspLeuHisArgIysValSerValGlnTyrCys 214
 |||||
 Db 649 GCCAGCTGCCACAGCAATGAGCCCTCGATCCCAAGATCTCGTCAGTACGCT 708
 QY 215 AspValLeuHisLysLysLys-----AlaLeuArgAsnLysLysLeuLeuArgAlaGlySer 232
 |||
 Db 709 GACGCGCTTCCACCCACGCGCTTTCGATGCGCGCCGCGCAAGGAGATG----- 756
 QY 233 GlyIlglyGlyGlyAspProGlyAspAlaPheGlyIleValAlaProPheAlaArg 252
 |||||
 Db 757 -----GACGCTATGCGATGCTGCGCTTCCGCGCGCGCG 792
 QY 253 ProSerValHisSerAspLeuMetIleArgGlyLysAlaSerAlaGlySerGlnAla 272
 |||||
 Db 793 CCCAGCTGCACAGTACCTCAAGTCAAGGCCAAGGTCCTTCGGGAAGGCCAGGCC 852
 QY 273 LysAspLysGluArgCysValIle 280
 |||||
 Db 853 CCGTAGAGGAGCAAGTGCACCATC 876
 RESULT 15
 ABL07789
 ID ABL07789 standard; cDNA; 1305 BP.
 AC ABL07789;
 XX 26-MAR-2002 (first entry)
 DT 26-MAR-2002 (first entry)
 XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 17849.
 DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 17849.
 XX Drosophila: developmental biology; cell signalling; insecticide;
 KW pharmaceutical; gene; ss.
 KM Drosophila melanogaster.
 OS Drosophila melanogaster.
 XX MO200117042-A2.
 PN 27-SEP-2001.
 XX 23-MAR-2001; 2001WO-US09231.
 PF 23-MAR-2001; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX (PEKE) PE CORP NY.
 PA Venter JC, Adams M, Li PWD, Myers EW;
 PI MPI: 2001-656860/75.
 DR P-PSDB; ABB63686.
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PR interactions -
 PT interactions -
 XX Claim 1: SEQ ID NO 17849; 21pp + Sequence Listing; English.
 PS The invention relates to an isolated nucleic acid detection reagent
 XX capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
 CC sequences (AB101840-AB16175) and the encoded proteins
 CC (AB57737-AB572072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pat_sequences.
 XX Sequence 1305 BP; 313 A; 390 C; 336 G; 266 T; 0 other;
 SQ Alignment Scores: 4.04e-61 Length: 1305
 Pred. No.:

Score: 615.00 Matches: 134
 Percent Similarity: 64.42% Conservative: 38
 Best Local Similarity: 50.19% Mismatches: 69
 Query Match: 42.44% Indels: 26
 DB: 23 Gaps: 6
 US-09-709-103-2 (1-281) x ABL07789 (1-1305)
 QY 18 SerIlePro--AlaLysAsnCysTyrArgMetValIleLeuGlySerSerLysValGly 36
 |||||
 Db 475 AGCGTGCATCGCGCAAAAGTCTGATCGATGCTGCGCTTCAATGCGATGCGCGCG 534
 QY 37 LysThrAlaIleValSerArgPheLeuThrGlyArgPheGluAspAlaTyrThrProThr 56
 |||||
 Db 535 AAGTCATCGATGCTGCGCGCTTCCGCGCAATGCGCTTCCGCGCGCGCGCGCGCGCG 594
 QY 57 IleGluAspPheHisArgLysPheTyrSerIleArgGlyValLysGluLeuAspIle 76
 |||||
 Db 595 ATCGAGAGTTCACCGCGCAATGATGATGCGATGCGAGTGAAGTCTTCAATGCGATATT 654
 QY 77 LeuAspThrSerGlyAsnHisProPheProAlaMetArgArgLeuSerIleLeuThrGly 96
 |||||
 Db 655 TTGATATCTCTCGCTATATCATCGCTTCCGCGCAATGCGATGCGATGCGATGCG 714
 QY 97 AspValPheIleLeuValPheSerLeuAspAsnArgAspSerPheGluGluValGlnArg 116
 |||||
 Db 715 GATCTCTTCATCTCGCTCTTCAGCATGATGATCCCGCGAGTCTTCGAGAGGTGTTCCG 774
 QY 117 LeuArgGlnGlnIleLeuAspThrLysSerCysLeuLysAsn-----Lys 131
 |||||
 Db 775 CTGCGGAGAACATCTCGGAGAACAGTGGCGTGCACCTAATCCGCGCTCCGCGTTTAA 834
 QY 132 ThrLysGluAsnValAspValProLeuValIleCysGlyAsnLysGlyAspArgAspPhe 151
 |||||
 Db 835 AAGAAGAGCTTCCAAAGATACCATGATATGCGCGGAATATGATGAGAGCTTT 894
 QY 152 TyrArgGluValAspGlnArgIleGluGlnLeuValGlyAspAspProGlnArgCys 171
 |||||
 Db 895 ---AAACTGTTCAGTGGAGGAGTGGGCTGATGCTGCGCGCGAGCAACTGCTGC 951
 QY 172 AlaTyrPheGluIleSerAlaLysAsnSerSerLeuAspGlnMetPheArgAlaLeu 191
 |||||
 Db 952 ACCTTTGTGAGTGTCTCGCTGCGTGCAGAAATTTACCGATGATGATCCACTGCTG 1011
 QY 192 PheAlaMetAlaLysLeuProSerGluMetSerProAspLeuHisArgLys--ValSer 210
 |||||
 Db 1012 TTACGCTCTCAATCTGCGCGTGGAGATGACCCCAATCATCATGCTGATGCTGCTCC 1071
 QY 211 ValGlnTyrCysAspValIleHisLysLysAlaLeuArgAsnLysLysLeuArgAla 230
 |||||
 Db 1072 GTTTTGGGGCGCGC-----TCGCCACTCCGCCACAC 1104
 QY 231 GlySerGlyGlyGlyGlyAspPro-----GlyAsp 241
 |||||
 Db 1105 GATCTCGGTGCGCGCGCGCGCAAGAAATGCCCTGCTCAAGAGAGATTACAGCAAT 1164
 QY 242 AlaPheGlyIleValAlaProPheAlaArgArgProSerValHisSerAspLeuMetTyr 261
 |||||
 Db 1165 GCGTGGGTGTGGAGCGCGCAAGCGGAGAGACCAAGATCGCGACGATGCTCATCTG 1224
 QY 262 IleArgGluLysAlaSerAla 268
 |||||
 Db 1225 ATGAGATCAAGACGATGGCG 1245

Search completed: December 28, 2002, 01:18:14
 Job time : 366 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 28, 2002, 01:13:57 ; Search time 87 Seconds

(without alignments)
1312.090 Million cell updates/sec

Title: US-09-709-103-2
Perfect score: 1449
Sequence: 1 MRLAMIKMKCPDSLSLIP.....IREKASAGQANDKRCVYS 281

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 363474 seqs, 203117208 residues

Total number of hits satisfying chosen parameters: 726948

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL-frame+pn.model -DRV-rlp
-Q/cgnt2_1/USPTO.spool/US09709103/runtat.20122002.162407.859/app_query.fasta.1.455
-DB-published_Applications_NA -QMT-fastap -SUFFIX-p2n.rmpb -MINMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX-blosum62
-TRANS-human40.ccl -LIST=45 -DOCALIGN=200 -THR SCORE=pcr -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODELOCAL -OUTFMT=ptc -NORR=ext -HEAPSIZE=500 -MITLEN=0
-MAXLEN=2000000000 -USER=US09709103 -ECGN_1.1.36 -runtat.20122002.162407.859
-NCPU=6 -ICPU=3 -NO_XLPRY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCGAPOP=6
-FCGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications_NA.*

- 1: /cgnt2_6/pdata/2/pubpna/US07_PUBCOMB.seq.*
- 2: /cgnt2_6/pdata/2/pubpna/PCIT_NEW_PUB.seq.*
- 3: /cgnt2_6/pdata/2/pubpna/US06_PUBCOMB.seq.*
- 4: /cgnt2_6/pdata/2/pubpna/US06_PUBCOMB.seq.*
- 5: /cgnt2_6/pdata/2/pubpna/US07_NEW_PUB.seq.*
- 6: /cgnt2_6/pdata/2/pubpna/PCITUS_PUBCOMB.seq.*
- 7: /cgnt2_6/pdata/2/pubpna/US08_NEW_PUB.seq.*
- 8: /cgnt2_6/pdata/2/pubpna/US08_PUBCOMB.seq.*
- 9: /cgnt2_6/pdata/2/pubpna/US09_NEW_PUB.seq.*
- 10: /cgnt2_6/pdata/2/pubpna/US09_PUBCOMB.seq.*
- 11: /cgnt2_6/pdata/2/pubpna/US10_NEW_PUB.seq.*
- 12: /cgnt2_6/pdata/2/pubpna/US10_PUBCOMB.seq.*
- 13: /cgnt2_6/pdata/2/pubpna/US60_NEW_PUB.seq.*
- 14: /cgnt2_6/pdata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	883	60.9	3082	10	US-09-778-963A-1
2	526	36.3	11221	10	US-09-778-963A-3
3	485	33.5	405	10	US-09-960-352-10273
4	370	25.5	368	10	US-09-864-761-21643

5	300	20.7	3346	9	US-09-764-868-67	Sequence 67, Appl
6	291	20.1	688	9	US-09-764-868-490	Sequence 490, App
7	283.5	19.6	2040	12	US-10-044-090-336	Sequence 336, App
8	278.5	19.2	197997	10	US-09-832-246-3	Sequence 3, Appl1
9	276.5	19.1	570	12	US-10-104-484-1	Sequence 1, Appl1
10	276.5	19.1	570	12	US-10-104-484-3	Sequence 3, Appl1
11	275.5	19.0	570	12	US-09-765-298A-25	Sequence 25, Appl
12	272.5	18.8	570	10	US-09-765-298A-27	Sequence 27, Appl
13	264.5	18.3	930	10	US-09-801-368-285	Sequence 285, App
14	259.5	17.9	536	10	US-09-867-701-5566	Sequence 5566, Ap
15	255.5	17.6	1271	12	US-10-044-090-110	Sequence 110, App
16	254	17.5	942	10	US-09-962-832-256	Sequence 256, App
17	249	17.2	847	10	US-09-822-949A-40	Sequence 40, Appl
18	245.5	16.9	952	10	US-09-801-368-287	Sequence 287, App
19	244	16.8	952	10	US-09-917-800A-1389	Sequence 1389, Ap
20	237	16.4	2418	9	US-09-764-868-84	Sequence 84, Appl
21	237	16.4	2479	9	US-09-764-868-69	Sequence 69, Appl
22	236.5	16.3	1085	9	US-10-067-813-3	Sequence 3, Appl1
23	236	16.3	1119	9	US-09-960-352-4253	Sequence 4253, Ap
24	231.5	16.0	452	10	US-09-917-800A-1616	Sequence 1616, Ap
25	226	15.6	1088	10	US-09-960-352-5830	Sequence 5830, Ap
26	225.5	15.6	454	10	US-09-960-352-5830	Sequence 111, App
27	225	15.5	1191	12	US-10-044-090-111	Sequence 983, App
28	214.5	14.8	1296	9	US-09-954-531-983	Sequence 1378, Ap
29	214.5	14.8	1296	9	US-09-954-531-1378	Sequence 6794, Ap
30	208	14.4	412	10	US-09-960-352-6794	Sequence 10, Appl
31	206	14.2	1199	9	US-09-794-257-10	Sequence 1, Appl1
32	203	14.0	1224	10	US-09-972-529-1	Sequence 1, Appl1
33	203	14.0	3192	10	US-09-788-654A-1	Sequence 3, Appl1
34	200.5	13.8	612	10	US-09-972-529-3	Sequence 91, Appl
35	194.5	13.4	836	10	US-09-770-445-674	Sequence 674, App
36	194	13.4	1274	10	US-09-925-302-91	Sequence 501, Appl
37	193	13.3	752	9	US-09-764-868-501	Sequence 44, Appl
38	193	13.3	1157	9	US-10-108-605-44	Sequence 6793, Ap
39	192	13.3	401	10	US-09-960-352-6793	Sequence 12, Appl
40	192	13.3	552	10	US-09-770-445-515	Sequence 515, Appl
41	190.5	13.1	884	10	US-09-919-580-42	Sequence 42, Appl
42	187	13.1	774	10	US-09-938-842A-420	Sequence 489, App
43	185.5	12.8	666	9	US-09-764-868-489	Sequence 2, Appl1
44	185.5	12.8	2734	9	US-10-091-613-2	
45	185.5	12.8	2964	9	US-10-091-613-2	

ALIGNMENTS

RESULT 1
US-09-778-963A-1
; Sequence 1, Application US/09778963A
; Patent No. US20020115172A1
GENERAL INFORMATION:
; APPLICANT: NEELAM, Beena et al
; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE
; FILE REFERENCE: C0001112
; CURRENT APPLICATION NUMBER: US/09/778, 963A
; CURRENT FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3082
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-778-963A-1

Alignment Scores:
Pred. No.: 6.21e-102
Score: 883.00
Percent Similarity: 77.78%
Best Local Similarity: 62.01%
Query Match: 60.94%
Length: 3082
Matches: 173
Conservative: 44
Mismatch: 46
Indels: 16
Gaps: 4

[illegible]

RESULT 7
US-10-044-090-336
; Sequence 336, Application US/10044090
; Patent No. US20020137081A1
; GENERAL INFORMATION:
; APPLICANT: Olga Bandman
; TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
; FILE REFERENCE: PA-0028 US
; CURRENT APPLICATION NUMBER: US/10/044,090
; CURRENT FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: PERL Program
; SEQ ID NO 336
; LENGTH: 2040
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID NO. US20020137081A1 3110662CB1
US-10-044-090-336

Alignment Scores:
Pred. No.: 4.14e-26 Length: 2040
Score: 283.50 Matches: 66
Percent Similarity: 55.21% Conservative: 40
Best Local Similarity: 34.388 Mismatches: 69
Query Match: 19.57% Indels: 17
DB: 12 Gaps: 3

US-09-709-103-2 (1-281) x US-10-044-090-336 (1-2040)

```
QY 25 TyrArgMetValIleLeuGlySerSerLysValGlyThrAlaIleValSerArgPhe 44
DB 202 TATAAGCTATGCTCTTGGCTGCGACGAGGCTTGGAAAGCTGCTTGCATGACATTTT 261
QY 45 LeuThrGlyArgPheGluAspAlaIleThrProThrIleGluAspPheHisArgLysPhe 64
DB 262 GTTCAAGGATTTTGTAGAAAATACGATCTTACGATGACAAATCTTATGAAAGCAA 321
QY 65 TyrSerIleArgGlyGluValTyrGlnLeuAspIleLeuAspThrSerGlyAsnHisPro 84
DB 322 GTTGAAGTAGATGACACACAGCTGATGCTGTAAGAACTGTGATCTGACAGAAAGCAA 381
QY 85 PheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIleLeuValPheSer 104
DB 382 TTACACGACATGAGGATTTATACATGAAAATGACAAAGATTTGCATTAGTTATTC 441
QY 105 LeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGlnGlnIleLeuAspThr 124
DB 442 ATCAGACACAGCTCCACATTTACGATTTACAGACCTGACAGACACAGATTTCTT 495
QY 125 LysSerCysLeuLysAsnLysThrLysGluAsnValAspValProLeuValIleCysGly 144
DB 496 -----CGAGTTAAAGACATGATGATGCTCCAAATGATTTGTTGGT 537
QY 145 AsnLysGlyAspArgAspPheTyrArgGluValAspGlnArgGluIleGlnIleVal 164
DB 538 AATTAAGTGAAGCTTGA-----GATGAAAGATTTGTAGGGAAGAACAA 582
QY 165 GlyAspAspProGlnArg-----CysAlaTyrPheGlnIleLeuSerAlaLysLys 180
DB 583 GGTCAAAATCTAGCAAGACAAATGGAACAACACTGCTTAAATCTTCTGCAAAATCA 642
QY 181 AsnSerLeuAspGlnMetPheArgAlaLeuPheAlaMetAlaLysLeuProSerGlu 200
DB 643 AAATAAATGTTAATGAGATCTTTATGACCTAGTCGCGCAATTAACAGAAAACCTCCA 702
QY 201 MetSerProAspLeuHisArgLysValSerValGln 212
DB 703 GTGCTGGGAAGGCTCGCAAAAAGTCATCATGTGAG 738
```

RESULT 8
US-09-822-246-3/c

; Sequence 3, Application US/09822246
; Patent No. US20020142383A1
; GENERAL INFORMATION:
; APPLICANT: MERKULOV, Gennady et al.
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; FILE REFERENCE: CLO01149
; CURRENT APPLICATION NUMBER: US/09/822,246
; CURRENT FILING DATE: 2001-04-02
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 19797
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)..(19797)
; OTHER INFORMATION: n = A,T,C or G
US-09-822-246-3

Alignment Scores:
Pred. No.: 2.01e-22 Length: 19797
Score: 278.50 Matches: 65
Percent Similarity: 55.21% Conservative: 41
Best Local Similarity: 33.85% Mismatches: 69
Query Match: 19.22% Indels: 17
DB: 10 Gaps: 3

US-09-709-103-2 (1-281) x US-09-822-246-3 (1-19797)

```
QY 25 TyrArgMetValIleLeuGlySerSerLysValGlyThrAlaIleValSerArgPhe 44
DB 43187 TATAAGTAGTGTGCTTGGCTGCGACGAGGCGTGGAAAGCTGCTTGCATGACATTTT 43128
QY 45 LeuThrGlyArgPheGluAspAlaIleThrProThrIleGluAspPheHisArgLysPhe 64
DB 43127 GTTCAAGGATTTTGTAGAAAATACGATCTTACGATGACAAATCTTATGAAAGCAA 43068
QY 65 TyrSerIleArgGlyGluValTyrGlnLeuAspIleLeuAspThrSerGlyAsnHisPro 84
DB 43067 GTTGAAGTAGATGACACACAGCTGATGCTTGAATCTTGCATGACGAAAGCAA 43008
QY 85 PheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIleLeuValPheSer 104
DB 43007 TTACACGACATGAGGATTTATACATGAAAATGACAAAGGCTTGCATTAGTTATTC 42948
QY 105 LeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGlnGlnIleLeuAspThr 124
DB 42947 ATCAGACACAGCTCCACATTTACGATTTACAGACCTGACAGACACAGATTTCTT 42894
QY 125 LysSerCysLeuLysAsnLysThrLysGluAsnValAspValProLeuValIleCysGly 144
DB 42893 -----CGAGTTAAAGACATGATGATGCTCCAAATGATTTGTTGGT 42852
QY 145 AsnLysGlyAspArgAspPheTyrArgGluValAspGlnArgGluIleGlnIleVal 164
DB 42851 AATTAAGTGAAGCTTGA-----GATGAAAGATTTGTAGGGAAGAACAA 42807
QY 165 GlyAspAspProGlnArg-----CysAlaTyrPheGlnIleLeuSerAlaLysLys 180
DB 42806 GGTCAAAATCTAGCAAGACAAATGGAACAACCTGCTTAAATCTTCTGCAAAATCA 42747
QY 181 AsnSerLeuAspGlnMetPheArgAlaLeuPheAlaMetAlaLysLeuProSerGlu 200
DB 42746 AAATAAATGTTAATGAGATCTTTATGACCTAGTCGCGCAATTAACAGAAAACCTCCA 42687
QY 201 MetSerProAspLeuHisArgLysValSerValGln 212
DB 42686 GTGCTGGGAAGGCTCGCAAAAAGTCATCATGTGAG 42651
```

RESULT 9

US-10-104-484-1
 ; Sequence 1, Application US/10104484
 ; Patent No. US20020150566A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Guan, Kun-Lian
 ; APPLICANT: Stewart, Scott
 ; TITLE OF INVENTION: METHOD OF INHIBITING CANCEROUS CELL PROLIFERATION USING RAS MUTAN
 ; TITLE OF INVENTION: GDP-BOUND CONFORMATION
 ; FILE REFERENCE: UMICH-0010
 ; CURRENT APPLICATION NUMBER: US/10/104,484
 ; CURRENT FILING DATE: 2002-03-25
 ; PRIOR APPLICATION NUMBER: 60/277,959
 ; PRIOR FILING DATE: 2001-03-23
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 1
 ; LENGTH: 570
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Nucleotide sequence of the GDP-bound Ras mutant, RasN17N69.
 ; US-10-104-484-1

Alignment Scores:

Pred. No.:	4,53e-26	Length:	570
Score:	276.50	Matches:	72
Percent Similarity:	52.74%	Conservative:	34
Best Local Similarity:	35.82%	Mismatches:	73
Query Match:	19.08%	Indels:	23
DB:	12	Gaps:	6

US-09-709-103-2 (1-281) x US-10-104-484-1 (1-570)

QY 25 TYRARGMETVALILELEUGLYSERSELYSVALGLYSTRHALLALEVALSERARGPHE 44
 DB 10 TATAAGCTGTGTGGTGGGCGCGCGGTGGGCAAGATGCGTACATCCACTG 69
 QY 45 LEUTHRGTYARGPHEGLUASPALATYRTHRPROTHRIEGLUASPPHEHISARGLYSPE 64
 DB 70 ATCCAGAACCACTTTGTGGAGCAATGACCCCACTATAGAGATTCCTCCGCAACAG 129
 QY 65 TYRSEIILEARGGLYGLUVALTYRGLINLEUASPILEUASPIRHSERGLYASNHISPRO 84
 DB 130 GTGTCATTGATGGGAGACGTGCTGTGGACATCTCTGATACCGCGGCGGAGAGAG 189
 QY 85 PHEPROALAMETARGTARGLEUSERILEUTHRGTYASPVALLPHEILEUVALPHESER 104
 DB 190 TACAGCGCATGCGGAACAGTACATGCGCACCGGGAGGCGTCTGTGTGTGTGCC 249
 QY 105 LEUASPRASPARGASERPHHEGLUGLVALGLNARGLEUARGGLNGLINLEUASPTH 124
 DB 250 ATCAACAAACCAAGCTCTTTGAGGACATCCACAGTACAGGAGCAATC----- 300
 QY 125 LYSSERCYSEULYSANLYSTHRLYSGIUSANVALASPVALLPROLEUVALILECYSGLY 144
 DB 301 -----AAACGGGTGAAGACTCGGATACCTGCCACAGTGGTGGG 345
 QY 145 ASNLYSGLYASPARGASRPHETARGGLUVALASPVALLASPVALLIEGLINLEUVAL 164
 DB 346 AACAAAGTGTACTG---GCTGACAGCACTGTGAATCTCGGAGGCTCAGGACTGCC 402
 QY 165 GLYASPRAPROGLNARGCYSAATYRPHHEGLUILESERALLALYSLSANSESRLEU 184
 DB 403 CGAAGC-----TACGGCATCCCTTACATCGAGACTCGGCCAAGACCCGCGAGAGTG 456
 QY 185 ASPGLINMERPHETARGALALEU-----PHEALAMETALLYSLEU--- 197
 DB 457 GAGGATGCTTTACAGCTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 516
 QY 198 ---PROSEGLINMERPHETARGALALEU-----PHEALAMETALLYSLEU--- 197
 DB 517 CTTCTGATGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 566

QY 217 LEU 217
 DB 567 CTG 569

RESULT 10

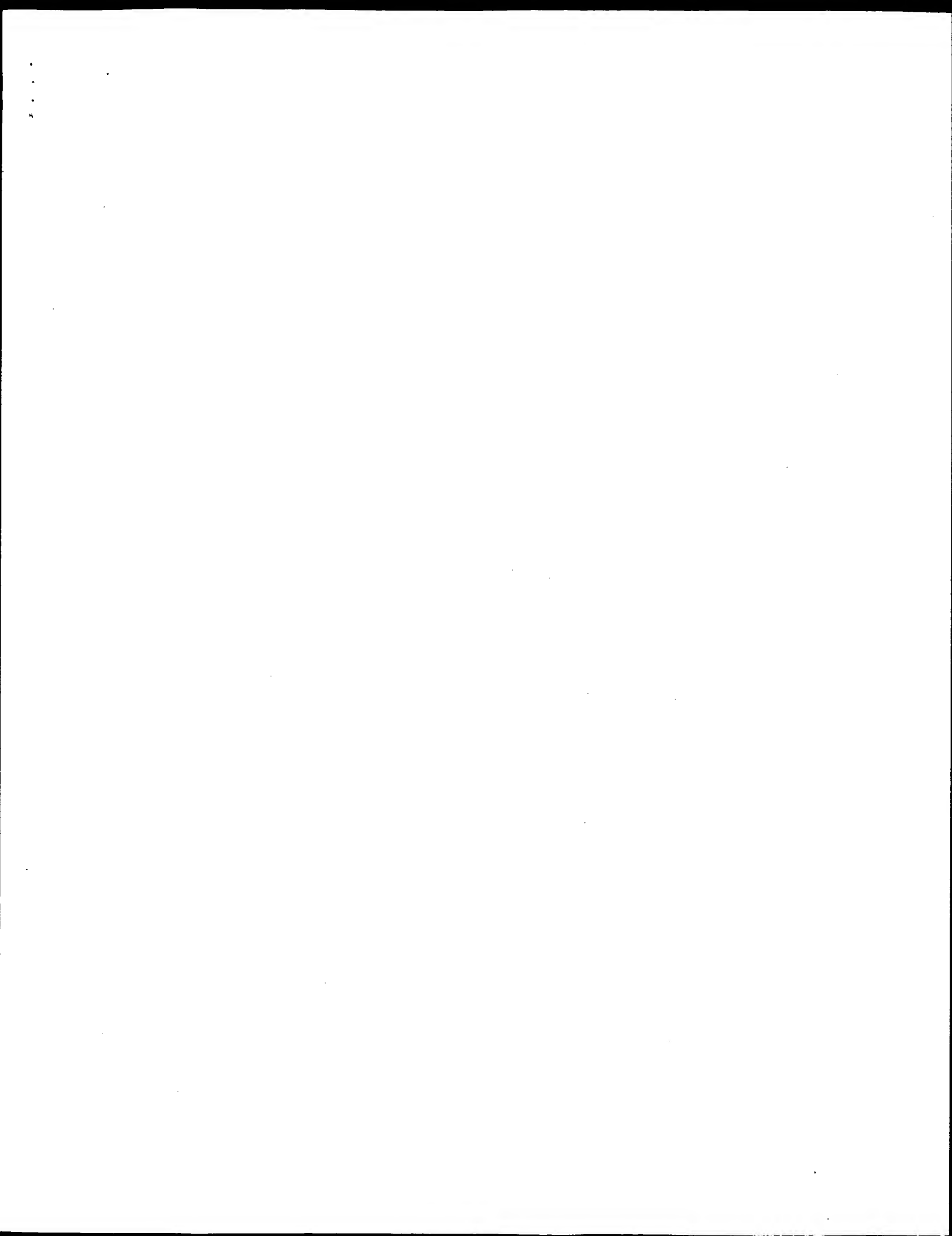
US-10-104-484-3
 ; Sequence 3, Application US/10104484
 ; Patent No. US20020150566A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Guan, Kun-Lian
 ; APPLICANT: Stewart, Scott
 ; TITLE OF INVENTION: METHOD OF INHIBITING CANCEROUS CELL PROLIFERATION USING RAS MU
 ; TITLE OF INVENTION: GDP-BOUND CONFORMATION
 ; FILE REFERENCE: UMICH-0010
 ; CURRENT APPLICATION NUMBER: US/10/104,484
 ; CURRENT FILING DATE: 2002-03-25
 ; PRIOR APPLICATION NUMBER: 60/277,959
 ; PRIOR FILING DATE: 2001-03-23
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 3
 ; LENGTH: 570
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Nucleotide sequence of the wild type Ras mutant, RasN17N69.
 ; US-10-104-484-3

Alignment Scores:

Pred. No.:	4,53e-26	Length:	570
Score:	276.50	Matches:	72
Percent Similarity:	52.74%	Conservative:	34
Best Local Similarity:	35.82%	Mismatches:	73
Query Match:	19.08%	Indels:	23
DB:	12	Gaps:	6

US-09-709-103-2 (1-281) x US-10-104-484-3 (1-570)

QY 25 TYRARGMETVALILELEUGLYSERSELYSVALGLYSTRHALLALEVALSERARGPHE 44
 DB 10 TATAAGCTGTGTGGTGGGCGCGCGGTGGGCAAGATGCGTACATCCACTG 69
 QY 45 LEUTHRGTYARGPHEGLUASPALATYRTHRPROTHRIEGLUASPPHEHISARGLYSPE 64
 DB 70 ATCCAGAACCACTTTGTGGAGCAATGACCCCACTATAGAGATTCCTCCGCAACAG 129
 QY 65 TYRSEIILEARGGLYGLUVALTYRGLINLEUASPILEUASPIRHSERGLYASNHISPRO 84
 DB 130 GTGTCATTGATGGGAGACGTGCTGTGGACATCTCTGATACCGCGGCGGAGAGAG 189
 QY 85 PHEPROALAMETARGTARGLEUSERILEUTHRGTYASPVALLPHEILEUVALPHESER 104
 DB 190 TACAGCGCATGCGGAACAGTACATGCGCACCGGGAGGCGTCTGTGTGTGTGCC 249
 QY 105 LEUASPRASPARGASERPHHEGLUGLVALGLNARGLEUARGGLNGLINLEUASPTH 124
 DB 250 ATCAACAAACCAAGCTCTTTGAGGACATCCACAGTACAGGAGCAATC----- 300
 QY 125 LYSSERCYSEULYSANLYSTHRLYSGIUSANVALASPVALLPROLEUVALILECYSGLY 144
 DB 301 -----AAACGGGTGAAGACTCGGATACCTGCCACAGTGGTGGG 345
 QY 145 ASNLYSGLYASPARGASRPHETARGGLUVALASPVALLASPVALLIEGLINLEUVAL 164
 DB 346 AACAAAGTGTACTG---GCTGACAGCACTGTGAATCTCGGAGGCTCAGGACTGCC 402
 QY 165 GLYASPRAPROGLNARGCYSAATYRPHHEGLUILESERALLALYSLSANSESRLEU 184
 DB 403 CGAAGC-----TACGGCATCCCTTACATCGAGACTCGGCCAAGACCCGCGAGAGTG 456
 QY 185 ASPGLINMERPHETARGALALEU-----PHEALAMETALLYSLEU--- 197
 DB 457 GAGGATGCTTTACAGCTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 516



CTA GCT GAT GAC GCA GCG CTC CTT GTC CTT GGC CTG GCT GCC GGC GCT GGC CTT
L A D D A A L L V L G L A A G A G L
. L M T Q R S L S L A W L P A L A F
S . . R S A P C P W P G C R R W P S

CTC GCG GAT GTA CAT GAG GTC GCT GTG TAC GCT GGG CCG GCG GCG GAA GGG TCC
81 108
L A D V H E V A V Y A G P A R E G C
S R M Y M R S L C T L G G R R A K G A
R G C T . G R C V R W A G A R R V P

CAC GAT GCC AAA GGC CTC GCC CAG GTC GCC GCC GCT GCC GGC
 H D A K G V A R V A A A A A A G P
 T M P K A S P G S P P P P L P A R
 R C Q R R R P G R R R R R C R P A

CAG CAG CTT CTT GTT CCG CAG CGC CTT CTT GTG CAG CAC GTC GCA GTA CTG CAC
 189
 Q Q L L L V P Q R L L L V Q H V A V L L
 S S F L F R S A F L C S T S Q Y C T P
 A A S C S A A P S C A A R R S T A P
 216

CGA GAC CTT GCG GTG CAG GTC TGG GCT CAT CTC GCT GGG CAG CTT GGC CAT GGC
243
R D L A V Q V W A H L A G Q L G H G
E T L R C R S G L I S L G S L A M A
R P C G A G L G S S R W A A W P W R
270

GAA GAG CGC GCG GAA CAT CTG GTC GCT GCT GTT CTT CTT GGC CGA GAT CAC
 E E R A E H L V Q A A V L L G R D L
 K S A R N I W S R L L F F L A E I S
 R A R G T S G P G C C S S W P R S R

GAA GTA GGC GCA GCC CTG GGG GTC GTC CAC CAG CTG CTC GAT CTC GCG GCG
 E V G A A L G V V A H Q L L D L A L
 K A Q R W G S S P T S C S I S R W
 S R R S A G G R R P P A A R S R A

GTC CAC CTC GCG GTA GAA GTC GCG ACC CTT GTT GCC GCA GAT GAC CAG GGG
 V H L A V E V A V T L V A A D D Q G
 S T S R . K S R S P L L P Q M T R G A
 P P R G R S R G H P C C R R . P G A

CAC GTC CAC GTT CTC GTT TTT GTT CTT GAG GCA AGA CTT GGT GTC GAG GAT
H V H V L L G F V L E A R L G V E D
T S T F S L V L F L R Q D L V S R I
R P R S P W F C S . G K T W C R G S

CTG CTG CCT GAG CCG CTG CAC CTC CTC GAA GGA GTC GCG GTT GTC CAG ACT GAA
L L P E P L H L L E G V A V V Q T E
C C L S R C T S S K E S R L S R L N
A A . A A A P P R R S R G C P D . T

CAC CAG GAT GAA AAC GTC TCC TGT TGG GAT GGA GAG GCG CCG CAT GGC GAG
 567 594
 H Q D E N V S C E D G E A P H G G E
 T R M K T S P V R M E R R M A G N

P G . K R L L . G W R G A A W R G T

CGG GTG GTT GCC GGA CGT GTC GAG GAT GTC GAG CTG GTA GAC CTC GCC GCG GAT
621
R V V A G R V E D V E L V D L A A D
G W L P D V S R M S S W . T S P R M
G G C R T C R G C R A G R P R R G W

GGA GTA GAA CTT GCG GTC GAA GTC GAT GGT AGG CGT GTA GGC CTC CTC GAA
 G V E L A V E V L D G R R V G V L E
 E . N L R W K S S M V G V . A S S K
 S R T C G G S P R W . A C R R P R S

GCG GCC GGT GAG GAA GCG CGA CAC GAT GGC CGT CTTT GCC CAC CTTT GGA CGA GCG
 A A G E E A R H D G R L A H L G R A
 R P V R K R D T M A V L P T L D E P
 G R . G S A T R W P S C P P W T S R

GAG GAT GAC CAT GCG ATA GCA GTT CTT GGC CGG GAT ACT CAG CTC CGA GTC GCT
E D D H A I A V L G R D T Q L R V A
R M T M R . Q F L A G I L S S E S L
G . P C D S S S W P G Y S A P S R S

CGG GCA CAT CTT GAT CAT CGC GGC CAG TTT CAT
R A H L L D H R G Q F H
G H I F L I I A A S F
G T S S . S S R P V S

OS-03-703-103-3

115-09-709-103-3 FEL

27

54

GGA ATT CCG AGC GGA GCC GGA GCC CCA AGC CCG AGC GGC CAG CCC GAG CAG

G I P S G A G A P S P S R A O P E O

AGC CCG CCA GCC GCT CAC CCC GCG TGC CAC CCC AGC AGC CCG CCG CTC TCTT
S P A A H P A C H P S D P O P L S

135
 GCC CTT CTC TCG GCC CCG CCG CCC CTC TGC CCA ATG AAA CTG
 A L L S A P R P S R P L C P M K L
 162

GCC GCG ATG ATC AAG AAG ATG TGC CCG AGC GAC TCG GAG CTG AGT ATC CCG GCC
 189
 A A M I K K M C P S D S E L S I P A
 216

243
AAG AAC TGC TAT CGC ATG GTC ATC CTC GGC TCG TCC AAG GTG GGC AAG ACG GCCC
K N C Y R M V I L G S S K V G K T A

SSRPVS
 15-09-709-103-3FL
 27
 54
 See 3, frame 1, 1-999
 nE

E L L S H G H P R L V Q G Q D G H
ATC GTG TCG CGC TTC CTC ACC GGC TTC GAG GAC GCC TAC ACG CCG ACC ATC 297 324
I V S R A F L T G R F E D A Y T P T I
S C R A S S P A S R T P T R L P S
R V A L P H R P L R G R L H A Y H R
GAG GAC TTC CAC CGC AAG TTC TAC TCC ATC CGC GGC GAG GTC TAC CAG CTC GAC 351 378
E D F H R K F Y S I R G E V Y Q L D
R T S T A S S T P S A A R S T S S T
G L P P Q V L L H P R R G L P A R H
ATC CTC GAC ACG TCC GGC AAC CAC CGC TTC CCC GCC ATG CGG CGC CTC TCC ATC 405 432
I L D T S G N H P P A M R R L S I
S S T R P A T T R S P P C G A A S P S
P R H V R Q P P V P R H A A P L H P
CTC ACA GGA GAT TTC ATC CTC GTG TTC AGT CTC GAC ACC AAG TCT TGC CTC TTC 459 486
L T G D V F I L V F S L D N R D S F
S Q E T F S S W C S V W T T A T P S
H R R R F H P G V Q S G Q P R L L S
GAG GAG GTG CAG CGC CTC AAG CAG CAG CTC GAC ACC AAG TCT TGC CTC TTC 513 540
E V Q R L R Q I L D T K S C L K
R R C S G S G S S T P S L A S R
G A A A Q A A D P R H Q V L L P Q E
AAC AAA ACC AAG GAG AAC GTG GAC GTG CCC CTC GTC ATC TGC GGC AAC AAG GGT 567 594
N K T K E N V D V P L V I C G N K G
T K P R R T W T C P W S S A A T R V
Q N Q G E R G R A P G H L R Q Q G
GAC CGC GAC TTC CAC CGC GAG GTG GAC CAG CGC GAG ATC GAG CAG CTC GTG GGC 621 648
D R D F Y R E V D Q R E I E Q L V G
T A T S T A R W T S A R S S S W W A
P R L L P R G G P A R D R A A G G R
GAC GAC CCC CAG CGC TGC GCC TAC TTC GAG ATC TCG GCC AAG AAG AAC AAG AGC 675 702
D D P Q R C A Y F E I S A K K N S S
T T P S A A P T S R S R P R R T A A
R P P A L R L L R D L G Q E E Q Q P
CTG GAC CAG ATG TTC CGC GGC CTC TTC GCC ATG GCC AAG CTC CCC AAG GAG ATG 729 756
L D Q M F R A L F A M A K L P S E M
W T R C S A R S S P W P S C P A R
G P D V P R A L R H G Q A A Q R D E
AGC CCA GAC CTC CAC CGC AAG GTC TCG GTG CAG TAC TGC GAC GTG CTC CAC AAG 783 810
S P D L L H R K V S V Q Y C D V L H K
A Q T C T A R S R C S T A T C T R
P R P A P Q G L G A V L R R A A Q E
AAG CGC CTG CGG AAC AAG AAG CTG CCG GGC GGC GGC GGC GGC GGC GGC 837 864
K A L R N K K L L R A G S G G G G G
R R C G T R S C C G P A A A A A A
G A A E Q E A A A G R Q R R R R
GAC CCG GGC GAC GCC TTT GGC ATC GTC GCA CCC TTT GGC CGC CGC CCC AGC GTC 891 918
D P G D A F G I V A P F A R R P S V
T R A T P L A S W H P S R A G P A Y
P G R R L W H R G T L R A P A Q R T
CAC AGC GAC CTC ATG ATC GTC CGC AAG GGC AGC GGC GGC GGC GGC GGC 945 972
H S D L M Y I R E K A S A G S Q A K
T A T S C T S A R R P A P A R P R
Q R P H V H P R E G Q R R Q P G Q
GAC AAG GAG CGC TGC GTC ATC AGC TAG GAG GAG CCC CGC GGC GGT GGC GAC ACA ACC 999 1026
D K E R C V I S A E P R R A G D T P
T R S A A S S A R S P A A L A T Q L
Q G A L R H Q L G A P P R W R H N L
TAA GGA GGA CCT TTT TGT TAA GTC AAA TCC AAC GGC CGG GTG CGC CCC AGG CCG 1053 1080
G G P F C V K S N G P V R P R P
K E D L F V K S N P T A R C A P G R
R R T F L L S Q I Q R P G A P Q A G
GGA CGC CGC GGC GAC TGC GGT CTC CCC TCC CGG CGA TCC GCC CCC AGC ACT GGG 1107 1134
G A R A D W R L P S R R S A P S T G
E A R T G V S P P G D P P P A L G
S A R G L A S P L P A I R P Q H W G
GAG CGC CCA CTG AAC GAA GGG ACG GTC ATC TGC TCC GGA AAG AAG AAG AAC 1161 1188
E A P L N R E G T V I C S G R K E N
R R H T E P R R D G H L L R K E R E
G A T E P R R D G H L L R K E R E
GGG CCA AGA CTG GGA GTC TTT CCC ACC CCC GGT CCC CCA TTG AAG CCC GCC AGC 1215 1242
G P R L G L F P T P P P P L R P A T
G Q D W D Y S P P P V P H G P P P
A K T G T I P H P R S P I E A R H P
CCC ATA ACT TTG GGA GGC AGG GCC CCG AAG GTG GAT TTA TCT TCT CAA AGA 1269 1296
P I T L G A R A Q P R V D L S S Q R D
H N F G S E G P A E G G F I F S K T
CCT AAG AGT GAG CGC GGC GTG GGC GGA TGT GAA GTT ATC CAG CCT CTG CTA 1323 1350
P K S E R G V G E G C E V I O P L L
L R V S A G W G R D V K L S S L C
E A A G G G G M S Y P A S A R
GGC TTC AAG AAA CCG TCA TGC CCG CTT GAG GGT CAG GAC CCA CGG GGC ATT ATC 1377 1404
G F K K P S C P L E G Q D P R G I I

A S R N R H A R L R V R T H G A L S
L Q E T V M P A . G S G P T G H Y L

1431
TTG TCT GTG ATT CCG GGT TGC TGT GAC ACC CGG TAG AGC CTC TGC CCT CCC GAA
L S V I P G C C D S R S L C P P E
C L . F R V A V T A G R A S A L P K
V C D S G L L . Q P V E P L P S R N

1485
ACT AAG CCG GGG GGC GTG GGT CAA ATC ATA GCC AAG TGA CTT GTT TAC ATG TGA
T K R G G V G Q I A K . L V Y M .
L S G G A W V K S . P S D L F T C E
A G G R G S N H S Q V T C L H V S

1539
GTG AAA CTG CAC AAA GGA ACA CAA AAC AAA ACT TGC ACT TTA ACG GTA GTT CCG
V K L H K G T Q N K T C T L T V V P
N C T K E T K T K L A L . R . F R
E T A Q R N T K Q N L H F N G S S G

1593
GTG TCA ACA TGG ACA CGA ACA AAA CCT TAC CCA GGT GTT TAT ACT GTG TGT GTG
V S T W T R T K P Y P G V T V C V
C Q H G H E Q N L T P Q V F I L C V C
V N M D T N K T L P R C L Y C V C V

1647
TGA GGT CTT TAA AGT TAT TGC TTT TTT TTT TAT AAT ATA CAA TAA AAT AAT
G L . S Y C F I W F F N I O . N N
E V F K V I A L F G F L I Y N K I I
R S L K K L L L Y L V F . Y T I K . F

1701
TTA AAA TGG AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA
L K W K K K K K K K K K K K K K K K
K N G K K K K K K K K K K K K K K K
K M E K K K K K K K K K K K K K K K

1755
AAA AAA AAA AAA AAA AAG CCG CTC GAG CAT GCA TCT AGA GGG CCG CAT
K K K K K K K K R P L E H A S R G P H
K K K K K K K S G R S S M H L E G R I
K K K K K K K A A R A C I . R A A A S

CAT GTA ATT AGT TAT GAA C
H V I S Y E
M . L V M N
C N . L

US-09-709-103-3' - numbering increases in 5' to 3' direction

27
GTT CAT AAC TAA TTA CAT GAT GCG GCC CTC TAG ATG CAT GCT CGA GCG GCC GGT
V H N . L H D A A L . M H A R A A A
F I T N Y M M R P S R C M L E R P L
S . L I T . C G P L D A C S S G R F

81
TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT
F F F F F F F F F F F F F F F F F
108

F F F F F F F F F F F F F F F F

135
TTT TTT TTT TCC ATT TTA AAT TAT TTT ATT GTA TAT TAA AAA ACC AAA TAA AGC
F F F S I L N Y F I V Y I K K T K . S
F F F P F I L I L Y C I L K N Q I K Q
F F F H F K L F Y C I L K N Q I K Q

189
AAT AAC TTT AAA GAC CTC ACA CAC AAT AAA CAG TAT AAA CAG TAT ANG GTT TTG
N N F K D L T H T Q Y K H L G K V L
I T L K T S H T H S I N T W V R F C
L . R P H T H T V . T P G . G F V

243
TTG GTG TCC ATG TTG ACA CCG GAA CTA CCG TTA AAG TGC AAG TTT TGT TTT GTG
F V S M L T P E L P L K C K F C F V
S C P C . H R N Y R . S A S F V L C
R V H V D T G T T V K V Q V L F C V

297
TTG CTT TGT GCA GTT TCA CTC ACA TGT AAA CAA CAA ACT TGG CTA TGA TTT GTG
F L C A V S L T C K Q V T W L . F D
S F V O F H S H V N K S L G Y D L T
P L C S F T H M . T S H L A M I . P

351
CCA CCG CCC GCT TAT TGA TGT CGG GAG GGC AGA GGC TCT ACC GGC TGT TGT CAC AGC
P R P P A . F R E G G R G S T G C H S
H A P P L L S F G R A E A L P A V T A
T P P R L L V S G G Q R L Y R L S Q Q

405
AAC CCG GAA TCA CAG ACA AGA TAA TGC CCC GTG GGT CCT GAC CCT CCA GCG GGC
N P E S O T R . C P V G P D P O A G
T R N H R Q D N A P W V L T L K R A
P G I T D K I M P R G S . P S S G H

459
ATG ACG GTT TCT TGA AGC CTA GCA GAG GGT GGA TAA CTT CAC ATC CCT CCC CCA
M T V S . S L A E A C . L H I P P P H
D G F L E A . Q R L D . N F T S L P T
R F L K P S R G W I T S H P S P T

513
CCC CCG GCT CAC TCT TAG CTC TTT GAG AAG ATA AAT CCA CCC TCG GCT GCG CCC
P R A H S . V F E K I N P P S A G P
P R S L L G L . E D K S T L G W A L

567
TCG CTC CCA AAG TTA TGG GGG TGG GCG GCC TCA ATG GGG GAC CCG GGG TGG GGA
S L P K L W G W R A S M G D R G G E
R S O S Y G G G P Q W G T G G G
A P K V M G V A G L N G G P G V G N

621
ATA GTC CCA GTC TTG GCC GGT TCT TCT TCC TTC CCG AGC AGA TGA TGA TCC TCC
I V P V L A R S L S F R S R . P S L
S Q S W P V L F P S G A D D T V P F
S P S L G P F S F L P E Q M T V P S

648

CTC GGT TGA GTG GCG CCT CCC CAG TGC TGG GGG CGG AGC ATC GCC GGG AGG GGA GAC 675 702
L G S V A P P Q C W G R I A G R G D
S V Q W R L P S A G G S P G G E T
R F S G A S P V L G A D R R E G R R
729 756
GCC AGT CCG GCG CTC CCG GCC TGG GGC GCA CCG GGC CGT TGG ATT TGA CTT
A S P R A L P A W G A P G R W I L
P V R A R S R P G A H R A V G F D L
Q S A R A P G L G R T G P L D L T
783 810
AAC AAA AAG GTC CTC CTT AGG TTG TGT CGC CAG CGC GGC GGC GGT CTT AGC TGA
N K K V L L L C R Q R G G A P S
T K R S S L G C V A S A A G L L A D
Q K G P P V V S P A R R G S L M
837 864
TGA CCG AGC GGT CCT TGT TGG CCT GGC TGC CCG GCG TGG CCT TCT CGC GGA
R S A L P C P W P G C R R W P S R G
D A A L L V L G L A A G L L A F S R M
T Q R S L S L A W L P A L A F S R M
891 918
TGT ACA TGA GGT CGC TGT GTA CCG TGG GCC GGC GGC CCA AGG GTG CCA CCA TGC
C T G R C V R W A G A R R V P R C
V H E V A V Y A G P A R E G G H D A
Y M R S L C T L G R R A K G A T M P
945 972
CAA AGC GGT CCG CCG GGT CGC CCG GGC GGC GGC TGC CCG CCC GCA GCA GGT
Q R R R P G R R R R R R C R P A A L
K G V A R V A A A A A G P A Q L
K A S P G S P P P P L P A R S S F
999 1026
TCT TGT TCC GCA GCG CCT TCT TGT GCA GCA CGT CGC AGT ACT GCA CCG AGA CCT
S C S A A P S C A A R R A S T A P R P
L V P Q R L L V Q H V A V L H R D L
L F R S A F L C S T S Q Y C T E T L
1053 1080
TGC GGT GCA GGT CTG GGC TCA TCT TGT TGG GCA GGT TGG CCA TGG CGA AGA GCG
C G A G L G S S R W A A W P W R R A
A V O V W A H L A G Q L G H G E R
R C R S G L I S L G S L A M A K S A
1107 1134
CGC GGA ACA TCT GGT CCA GGC TGC TGT TCT TGG CCG AGA TCT CGA AGT AGG
R G T S G P G C C S S W P R S R S R
A E H L V Q A A V L L G R D L E V G
R N I W S R L L F F L A E I S K A
1161 1188
CGC AGC GGT GGG GGT GGT CCG CCA GCT GCT CGA TCT CGC GGT GGT CCA CCG
R S A L G G R R P A A A R S R A G P P
Q A L G V V A H Q L L D L A L V H L
Q R W G S S P T S C S I S R W S T S
1215 1242
CGC GGT AGA AGT CCG GGT CAC CCT TGC CCG AGA TGA CCA GGG GCA GGT CCA
R G R S R G H P C C R R P G A R P

A V E V A V T L V A A D D Q G H V H
R K S R S P L L P Q M T R G T S T
1269 1296
CGT TCT CCT TGG TTT TGT TCT TGA GGC AAG ACT TGG TGT CGA GGA TCT GCT GCC
R S L P W F C S G K T W C R G S A A P
V L L G F V L E A R L G V E D L A
F S L V L F L R Q D L V S R I C C L
1323 1350
TGA GCC GCT GCA CCT CCT CGA AGG AGT CCG GGT TGT CCA GAC TGA ACA CCA CCA
A A A P P R S R G C P D T P G
E P L H L L E G V A V V Q T E H Q D
S R C T S S K E S R L S R L N T R M
1377 1404
TGA AAA CGT CTC CTG TGA GGA TGG AGA GGC GCC GCA TGG CCG GGA AGC GGT GGT
K R L L G W R G A A W R G T G V
E N V S C E D E A P H G G E R V
K T S P V R M E R R R M A G N G W L
1431 1458
TGC CCG AGC TGT CGA GGA TGT CGA GGT GGT AGA CCT CGC GGC GGA TGG AGT AGA
C R T C R G C R A G P R R G W S R
A G R V E D V E L V D L A A D G V E
P D V S R M S S W T S P R M E N
1485 1512
ACT TGC GGT GGA AGT CCT CGA TGG TAG GCG TGT AGG GGT CCT CGA AGC GGC CCG
T C G G S P R W A C R R P R S G R
L A V E V L D G R R V G V L E A A G
L R W K S S M V G V A S S K R P V
1539 1566
TGA GGA AGC GCG ACA CCA TGG CCG TCT TGC CCA CCT TGG AGC AGC GGA TGA
G S A T R W P S C P P W T S R G
E A R H D G R L A H L G R A E D D
R K R D T M A V L P T L D E P R M T
1593 1620
CCA TGC GAT AGC AGT TCT TGG CCG GGA TAC TCA GCT CCG AGT CCG TGG GGC AGA
P C D S S S W P G Y S A P S R S G T
H A I A V L G R D T Q L R V A R H I
M R Q F L A G I L S S E S L G H I
1647 1674
TCT TCT TGA TCA TCG CCG CCA GTT TCA TTG GGC AGA GGC GCC GCG AGG GCG GGC
S S S S R P V S L G R G A A R A G
L L D H R G Q F H W A E G P R G R A
F L I I A A S F I G Q R G R E G R
1701 1728
GCG GGG CCG AGA GAA GGG CAG AGA GCG GGT GAG GGT TGG GGT GGC AGC CCG
A G P R E G Q R A A E G R W G G T R
R G R E K G R E R L R V A G V A R G
G A E R R A E S G G S L G W H A G
1755 1782
GGT GAG CCG CTG GAG GGC TCT GCT CCG GGT GGC GGC TCG GGC TTG GGC CTC
G E R L E G S A R A G R G S G L G L
V S G W R A L L G L G A A R A W G S
A A G G L C S G W A R L G L G A P

CGG CTC CGC TCG GAA TTC C
R L R S E F
G S A R N S
A P L G I

TRANSLATE [Six-Frame, One-Letter]: q

SEQ: q

No changes made since last save, file(s) not written.

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: December 30, 2002, 14:39:31 ; Search time 3141 Seconds
(without alignments)
2612.859 Million cell updates/sec

Title: US-09-709-103-1f1

Perfect score: 1450

Sequence: 1 MCLAMIKMCPDSELSIP.....REKASAGSQAKDKERCVIS. 282

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2054640 segs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xlp
-Q=/cgn2.1/USPTO.spool/TRANSUS09709103/runat.30122002.143925.20652/APP.query.fasta.1.455
-DB=genEmbl -QPRF=fastap -SUFFIX=p2n.rge -MINMATCH=0 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human4.0.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pco -NOM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=TRANSUS09709103 @CGN 1.1 3637 @runat.30122002.143925.20652 -NCPU=6
-ICPU=3 -NO XLPXY -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:*
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_in:*
5: gb_ov:*
6: gb_pac:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: gb_vl:*
16: em_ba:*
17: em_fun:*
18: em_hum:*
19: em_hum:*
20: em_mu:*
21: em_mu:*
22: em_ov:*
23: em_ov:*
24: em_pac:*
25: em_ph:*
26: em_pl:*
27: em_ro:*
28: em_sts:*
29: em_un:*

29: em_vl:*
30: em_hcg_hum:*
31: em_hcg_inv:*
32: em_hcg_other:*
33: em_hcg_mus:*
34: em_hcg_pln:*
35: em_hcg_rtd:*
36: em_hcg_mam:*
37: em_hcg_vrt:*
38: em_sy:*
39: em_hgo_hum:*
40: em_hgo_mus:*
41: em_hgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1449	99.9	846	9	AF498923 Homo sapi
2	1449	99.9	979	9	AF172846 Homo sapi
3	1449	99.9	1187	9	AF153192 Homo sapi
4	1449	99.9	1740	9	AF069506 Homo sapi
5	1449	99.9	1758	9	BC018041 Homo sapi
6	1436	99.0	1746	9	AF177335 Homo sapi
7	1405.5	96.9	1612	10	BC034166 Mus muscu
8	1405.5	96.9	1623	10	AF009246 Mus muscu
9	1399.5	96.5	1616	10	AF239157 Rattus no
10	1387	95.7	1833	9	AF222979 Homo sapi
11	1387	95.7	1833	9	AF222979 Homo sapi
12	1384	95.4	5141	9	AF262018 Homo sapi
13	1370	94.5	1791	10	AF603710 Mouse DNA
14	1370	94.5	2478	2	AC025909 Mus muscu
15	1356	93.4	1835	2	AC090608 Homo sapi
16	1355	93.4	1835	2	AC073621 Homo sapi
17	985	67.9	1625	2	AC122995 Rattus no
18	893	61.6	2832	6	AX393244 Sequence
19	893	61.6	2832	6	AF279143 Homo sapi
20	893	61.6	2973	6	AX393267 Sequence
21	867	59.8	3020	6	BC013419 Homo sapi
22	867	59.8	3020	6	AX393362 Sequence
23	790	54.5	2629	9	AF134409 Rattus no
24	790	54.5	2629	9	AL831936 Homo sapi
25	526	36.3	1147	10	AL022334 Human DNA
26	512	35.3	2110	10	AC076974 Mus muscu
27	359.5	24.8	1815	2	AC010562 Drosophi
28	359.5	24.8	2076	2	AC018039 Drosophi
29	359.5	24.8	2854	3	AE003560 Drosophi
30	359	24.8	2294	10	BC026377 Mus muscu
31	342	23.6	4412	9	AB062937 Macaca fa
32	334	23.0	1875	9	BC008065 Homo sapi
33	334	23.0	2505	9	AB076889 Homo sapi
34	334	23.0	4167	6	AX477382 Sequence
35	334	23.0	1432	9	AL353619 Human DNA
36	334	23.0	1438	9	AC021583 Homo sapi
37	334	22.5	1996	9	AC021582 Homo sapi
38	326	22.5	597	9	AY056037 Homo sapi
39	326	22.5	774	9	AY059641 Homo sapi
40	326	22.5	2827	9	AB076888 Homo sapi
41	326	22.5	3391	9	AK096500 Homo sapi
42	326	22.5	1754	9	BC030660 Homo sapi
43	324	22.3	1269	2	AC006538 Homo sapi
44	324	22.3	1269	2	AC091518 Mus muscu
45	324	22.3	2013	2	AC073816 Mus muscu

ALIGNMENTS

RESULT 1

AF498923 AF498923 846 bp mRNA linear PRI 01-MAY-2002
 LOCUS Homo sapiens activator of G protein signaling (RASD1) mRNA,
 complete cds.
 ACCSSON AF498923.1 GI:20379021
 VERSION AF498923.1
 KEYWORDS
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 846)
 AUTHORS Pohl, H.L., Ikeda, S.R. and Aronstam, R.S.
 TITLE Direct Submission
 JOURNAL Submitted (05-APR-2002) cDNA Resource Center, Guthrie Research
 Institute, One Guthrie Square, Sayre, PA 18840, USA
 FEATURES
 source
 1. .846
 /location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /tissue_type="brain"
 1. .846
 /gene="RASD1"
 /note="AGS1"
 1. .846
 /genes="RASD1"
 /codon_start=1
 /product="activator of G protein signaling"
 /protein_id="AA021071.1"
 /db_xref="GI:20379022"
 /translations="MKLAAMIKMCPDSSELSIPAKNVCYRMVILGSSKVGKTAIVSRF
 LTGRFDAYPTIEDPRKFSYIRGEVQIDLDITSGNHPFPAMRRLSLITGDFVILV
 FSDNRDSFEVQRLRQQLIDTKSLKNKTENVDVPLVCGNKGDROFREVDDREI
 EQLVDDPQRCAYFEISAKNSLDQMFALFAMAKLPSEMSFDLHRKVSQVCVLH
 KQALRNKLLLRAGSGGGGDPDAFIVAPFARRPSVHSDMLYIREKASASQAKDKE
 RCVIS"
 BASE COUNT 176 a 288 c 252 g 130 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,42e-127 Length: 846
 Score: 1449.00 Matches: 281
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 99.93% Indels: 0
 DB: 9 Gaps: 0
 US-09-709-103-1F1 (1-282) x AF498923 (1-846)
 QY 1 MetLysLeuAlaAlaMetIleLysLysMetCysProSerAspSerGluLeuSerIlePro 20
 Db 1 ATGAACTGGCGCGATGATCAAGAGATGTGCCGAGCGACTCGGAGCTGAGTATCCCG 60
 QY 21 AlalysAsnCysTyrArgMetValIleLeuGlySerSerLysValGlyLysThrAlaIle 40
 Db 61 GCCAAGAACTGCTATCGATGTCATCTCGGCTCGTCCAAAGTGGCAAGACGGCCATC 120
 QY 41 ValSerArgPheLeuThrGlyArgPheGluAspAlaTyrThrProThrIleGluAspPhe 60
 Db 121 GTGTGCGCGTCTCTCCACCGCGCGCTTCGAGGAGCGCTTACACGCTACCATCGAGGACTTC 180
 QY 61 HisArgLysPheTyrSerIleArgGlyGluValTyrGlnLeuAspIleLeuAspThrSer 80
 Db 181 CACGCGAAGTTCTACTCATCCCGCGCGAGGTCTACAGCTCGACATCTCTCGACGTC 240
 QY 81 GlyAsnHisProPheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIle 100
 Db 241 GGCAACACCGTTCCTCCCGCGCATCGCGCGCTCTCCATCTCCACAGAGACGTTTTCATC 300
 QY 101 LeuValPheSerLeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGlnGln 120
 Db 301 CTGGTGTTCAGTCTGGCAACACCGCGACTCTCTCGAGGAGGTGAGCGGCTCAGCGACG 360

QY 121 IleLeuAspThrLysSerCysLeuLysAsnLysThrLysGluAsnValAspValProLeu 140
 Db 361 ATCTCTCGACACCAAGTCTTGCTCAAGAACAAACCAAGGAGACGTGGAGCTGCCCTG 420
 QY 141 VallieCysGlyAsnLysGlyAspArgAspPheTyrArgGluValAspGlnArgGluIle 160
 Db 421 GTCATCTCGCGCACACAGGGTGCACCGGACTTCTACCGGAGGTGGACCAAGCGAGATC 480
 QY 161 GluGlnLeuValGlyAspAspProGlnArgCysAlaTyrPheGluIleSerAlaLysLys 180
 Db 481 GAGCAGCTGGTGGGCGAGCAGCCCCAGCGCTGCGCTACTTCGAGATCTCGGCCAAGAAG 540
 QY 181 AsnSerSerLeuAspGlnMetPheArgAlaLeuPheAlaMetAlaLysLeuProSerGlu 200
 Db 541 AACAGCAGCTGGACCATGTTCCGCGCGCTCTTCGCCATGGCCAGCTGCCACGCGAG 600
 QY 201 MetSerProAspLeuHisArgLysValSerValGlnTyrCysAspValLeuHisLysLys 220
 Db 601 ATGAGCCAGACCTGCACCGCAAGGTCTCGGTGCAGTACTCGGACGTGCTGCACAAGAAG 660
 QY 221 AlaLeuArgAsnLysLysLeuLeuArgAlaGlySerGlyGlyGlyGlyGlyAspProGly 240
 Db 661 GCGCTCGGAAACAAGAGCTGTCGCGCGCGCAGCGCGCGCGCGCGCGCGCGCGCG 720
 QY 241 AspAlaPheGlyIleValAlaProPheAlaArgArgProSerValHisSerAspLeuMet 260
 Db 721 GACGCTTTGGCATCTGGCACCCTTCGCGCGCGCGCGCGCGCGCTACACAGGACTCATG 780
 QY 261 TyrIleArgGluLysAlaSerAlaGlySerGlnAlaLysAspLysGluArgCysValIle 280
 Db 781 TACATCCGCGAAGCG 840
 QY 281 Ser 281
 Db 841 AGC 843
 RESULT 2
 LOCUS AF172846 Homo sapiens ras-related protein (DEXRAS1) mRNA, complete cds.
 DEFINITION AF172846
 ACCESSION AF172846
 VERSION AF172846.1 GI:6014488
 KEYWORDS
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 979)
 AUTHORS Tu, Y. and Wu, C.
 TITLE Cloning, expression and characterization of a novel human
 Ras-related protein that is regulated by glucocorticoid hormone
 JOURNAL Biochim. Biophys. Acta 1489 (2-3), 452-456 (1999)
 MEDLINE 20135605
 PUBMED 10673050
 REFERENCE 2 (bases 1 to 979)
 AUTHORS Tu, Y. and Wu, C.
 TITLE Direct Submission
 JOURNAL Submitted (27-JUL-1999) Department of Cell Biology, University of
 Alabama at Birmingham, 1670 University Blvd., Birmingham, AL 35294,
 USA
 FEATURES
 Location/Qualifiers
 1. .979
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 1. .979
 /gene="DEXRAS1"
 64. .909
 /gene="DEXRAS1"
 /note="similar to Mus musculus ras-related protein encoded
 by GenBank Accession Number AF009246; regulated by
 glucocorticoid hormone"
 /codon_start=1
 /product="ras-related protein"

```

/protein_id="AA01364.1"
/db_xref="GI:6014489"
/translation="MKLAAMIKKMPDSELSIPAKNCYRWVILGSSKVKTAIVSRF
LTGRFEDAVPTIEDPHRRKFSYIRGEVQDILDTSGNHPFPMRRLSLITGVFLLV
FSLNDRDSFEVQRLROOILDTKSLKTKTKTENVDPLVICGNGDDDFXEVNOREI
EQLVGDPOKCAVEISAKNSLSDMPALFMAKLPSESPDLHRKVSQYCDVILH
KKALRNKTLIRAGSGGGGDPDAFGIYAPFARPSVHSDLMYIRKASAGSQAQKE
RCVTS"
BASE COUNT      195 a      347 c      284 g      153 t
ORIGIN
Alignment Scores:
Pred. No.:      1,69e-127      Length:      979
Score:          1449.00      Matches:      281
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    99.93%      Indels:      0
DB:             Gaps:      0
US-09-709-103-1f1 (1-282) x AF172846 (1-979)
QY      1 MetLysLeuAlaAlaMetLysLysMetCysProSerAspSerGluLeuSerIlePro 20
DB      64 ATGAAGACTGCGCGGATGATCAAGAAAGATGTGCCGAGGACTCGAGCTGAGTATCCCG 123
QY      21 AlAlYsAsnCysEtyrArGmetValIleleuGlYserSerLysValGlyThrAlaIle 40
DB      124 GCCAAGACTGCTATGTCATGTCATCTCGCTCTCCAGGTGGCAAGACGGCCATC 183
QY      41 ValSerArgPheLeuThrGlyArgPheGluAspAlaTyrrThrProThrIleGluAspPhe 60
DB      184 GTGTCCGCTTCTCTCAACCGCGCTTCGAGAGCGCTTACACCGCTTACATCGAGCATTC 243
QY      61 HisArgLysPheThrSerIleArgGlyGluValTyrrGluLeuAspIleLeuAspThrSer 80
DB      244 CACCGAAGTCTTACTTCATCCCGCGGAGGTCTACAGCTGCAATCCTCGACACGTC 303
QY      81 GlyAsnHisproPheProAlaMetArgArgLysSerIleLeuThrGlyAspValPheIle 100
DB      304 GGCAACCAACCCGTTCCCGCGCATCGCGGCTCTCCATCTCCACAGACACGTTTTCATC 363
QY      101 LeuValPheSerLeuAspAsnArgAspSerPheGluGluValGlnArgLysGlnGln 120
DB      364 CTGTGTTCAGTCTGGAACAACCGCACTCTTCGAGAGGTGACGCGCTCAAGCAGCAG 423
QY      121 IleLeuAspThrLysSerCysLeuLysAsnLysThrLysGluAsnValAspValProLeu 140
DB      424 ATCTCGACACCAAGTCTTGCCTCAAGAAACAACAGAGAAAGTGAACGTGACGCTCG 483
QY      141 ValIleCysGlyAsnLysGlyAspArgAspPheTyrrArgGluValAspGlnIle 160
DB      484 GTATCTGCGGCAACAAAGGTGACCGGACTTTCACCGGAGGTGACCAAGCCGAGATC 543
QY      161 GluGluLeuValGlyAspAspProGlnArgCysAlaLysPheGluIleSerAlaLysLys 180
DB      544 GACGAGCTGTGGGCGACGACCCGCGCTGCGCTTCCAGATCTCGGCGCAAGAG 603
QY      181 AsnSerSerLeuAspGlnMetPheArgAlaLeuPheAlaMetLysLysLeuProSerGln 200
DB      604 AACAGCAGCTGACGAGATGTTCCGCGCTCTTCGCCCAATGCGCAAGCGCCACGAG 663
QY      201 MetSerProAspLeuHisArgLysValSerValGlnTyrrCysAspValLeuHisLysLys 220
DB      664 ATGAGGCCAATCCTGCAACCGCAAGTCTCGGTGCAAGTACGCACTGCTGCAACAAGAG 723
QY      221 AlAlYsAsnCysEtyrArGmetValIleleuGlYserSerLysValGlyThrAlaIle 40
DB      724 GCGCTGCGGAACAAGAGCTGTGCGGCGGCGACGCGCGCGCGCGCGCGCGCGCG 783
QY      241 AspAlaPheGlyIleValAlaProPheAlaArgArgProSerValHisSerAspLeuMet 260
DB      784 GACGCTTTGGCATGTGGCACTCTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 843

```

```

QY      261 TyrIleArgGluLysAlaSerAlaGlySerGlnAlaLysAspLysGluArgCysValIle 280
DB      844 TACATCCGCGAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 903
QY      281 Ser 281
DB      904 AGC 906
RESULT 3
AF153192      1187 bp      mRNA      linear      PRI 02-JUN-1999
LOCUS
DEFINITION
Homo sapiens ras-related protein mRNA, complete cds.
ACCESSION
AF153192
VERSION
AF153192.1 GI:4960166
KEYWORDS
SOURCE
ORGANISM
Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 1187)
AUTHORS
Kempainen, R.J.
TITLE
Identification of human pituitary Dextral
JOURNAL
Unpublished
2 (bases 1 to 1187)
AUTHORS
Kempainen, R.J.
TITLE
Direct Submission
JOURNAL
Submitted (20-MAY-1999) Anatomy & Physiology, Auburn University,
College of Veterinary Medicine, Auburn, AL 36849, USA
FEATURES
Source
1..1187
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="2"
/map="2q32"
/tissue_type="pituitary gland"
1..846
/note="Dextral; similar to Mus musculus ras-related
protein encoded by Genbank Accession Number AF09246"
/codon_start=1
/product="ras-related protein"
/protein_id="A034621.1"
/db_xref="GI:4960167"
/translation="MKLAAMIKKMPDSELSIPAKNCYRWVILGSSKVKTAIVSRF
LTGRFEDAVPTIEDPHRRKFSYIRGEVQDILDTSGNHPFPMRRLSLITGVFLLV
FSLNDRDSFEVQRLROOILDTKSLKTKTKTENVDPLVICGNGDDDFXEVNOREI
EQLVGDPOKCAVEISAKNSLSDMPALFMAKLPSESPDLHRKVSQYCDVILH
KKALRNKTLIRAGSGGGGDPDAFGIYAPFARPSVHSDLMYIRKASAGSQAQKE
RCVTS"
BASE COUNT      249 a      396 c      362 g      180 t
ORIGIN
Alignment Scores:
Pred. No.:      2.13e-127      Length:      1187
Score:          1449.00      Matches:      281
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    99.93%      Indels:      0
DB:             Gaps:      0
US-09-709-103-1f1 (1-282) x AF153192 (1-1187)
QY      1 MetLysLeuAlaAlaMetLysLysMetCysProSerAspSerGluLeuSerIlePro 20
DB      1 ATGAAGACTGCGCGGATGATCAAGAAAGATGTGCCGAGGACTCGAGCTGAGTATCCCG 60
QY      21 AlAlYsAsnCysEtyrArGmetValIleleuGlYserSerLysValGlyThrAlaIle 40
DB      61 GCCAAGACTGCTATGTCATGTCATCTCGCTCTCCAGGTGGCAAGACGGCCATC 120
QY      41 ValSerArgPheLeuThrGlyArgPheGluAspAlaTyrrThrProThrIleGluAspPhe 60
DB      121 GTGTCCGCTTCTCTCAACCGCGCTTCGAGAGCGCTTACACCGCTTACATCGAGCATTC 180

```


Qy 61 HisArgLysPheTyrSerIleArgGlyGluValTyrGlnLeuAspIleLeuAspThrSer 80
 Db 181 CACCGCAAGTTCTACTCCATCGCGGCGAGGTCTACGAGCTCGACATCTCTCGACACGTCC 240
 Qy 81 GlyAsnHisProPheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIle 100
 Db 241 GGCAACACACCGTTCCCGCCATCGCGGCGCTCTCCATCTCTACAGAGAGCTTTTCATC 300
 Qy 101 LeuValPheSerLeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGlnGln 120
 Db 301 CTGTGTGTTTCAGCTTGGACAAACCGGACCTCTTCGAGGAGGTGACGCGCTCAGGACGAG 360
 Qy 121 IleLeuAspThrLysSerCysLeuLysAsnLysThrLysGluAsnValAspValProLeu 140
 Db 361 ATCTCTGACACCAAGTCTTGCTCTCAAGAACAAACCAAGGAGAGAGCTGCGCCCTG 420
 Qy 141 ValIleCysGlyAsnLysGlyAspArgAspPheTyrArgGluValAspGlnArgGluIle 160
 Db 421 GTCATCTCGCGGCACAAAGGGGACCGGACCTCTTACCGGAGGTGGACCGCGGAGATC 480
 Qy 161 GluGlnLeuValGlyAspAspProGlnArgCysAlaTyrPheGluIleSerAlaLysLys 180
 Db 481 GAGCAGCTGTGTGGCGACGACCCCGCGCTGCGCTACTTCGAGATCTCGGCCAAGAAG 540
 Qy 181 AsnSerSerLeuAspGlnMetPheArgAlaLeuPheAlaMetAlaLysLeuProSerGlu 200
 Db 541 AACAGCAGCTGGACCAAGATCTTCGCGCGCTCTTCGCCATGCGCAAGCTGCCGACGAG 600
 Qy 201 MetSerProAspLeuHisArgLysValSerValGlnTyrCysAspValLeuHisLysLys 220
 Db 601 ATGAGCCAGACCTGACCGCAAGGTCTCGGTGAGTACTGCGACGTGCTGCACAGAAG 660
 Qy 221 AlaLeuArgAsnLysLysLeuLeuArgAlaGlySerGlyGlyGlyGlyAspProGly 240
 Db 661 CGCTGCGGAACAAGAGCTGCTGCGGCGCGGACGCGCGCGCGCGCGCGCGCGCGG 720
 Qy 241 AspalapheGlyIleValAlaProPheAlaArgArgProSerValHisSerAspLeuMet 260
 Db 721 GACGCTTGTGCGATCGTGGACCTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 780
 Qy 261 TyrIleArgGluLysAlaSerAlaGlySerGlnAlaLysAspLysGluArgCysValIle 280
 Db 781 TACATCCGCGAGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 840
 Qy 281 Ser 281
 Db 841 AGC 843

RESULT 4
 AF069506 1740 bp mRNA linear PRI 10-JAN-2000
 LOCUS Homo sapiens activator of G protein signaling (AGS1) mRNA, complete cds.
 DEFINITION
 ACCESSION AF069506
 VERSION AF069506.1 GI:4959037
 KEYWORDS
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1740)
 Cismowski,M.J., Takesono,A., Ma,C., Lizano,J.S., Xie,X.,
 Fuernkranz,H., Lanier,S.M. and Duzic,E.
 Genetic screens in yeast to identify mammalian nonreceptor
 modulators of G-protein signaling
 Nat. Biotechnol. 17 (9), 878-883 (1999)
 JOURNAL 99403338
 MEDLINE 10471929
 PUBMED 2 (bases 1 to 1740)
 Cismowski,M.J., Fuernkranz,H., Ma,C., Spruyt,M., Xie,X.,
 Lanier,S.M. and Duzic,E.
 Direct Submission
 JOURNAL Submitted (01-JUN-1998) Receptor Pharmacology/Biochemistry, Cadus

Pharmaceutical Corporation, 777 Old Saw Mill River Rd., Tarrytown,
 NY 10591, USA
 FEATURES
 source
 1. 1740

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /tissue_type="liver"
 /dev_stages="adult"

gene

1. 1740

/gene="AGS1"

146 991

/gene="AGS1"

/note="ras-related G protein; similar to Mus musculus
 product encoded by GenBank Accession Number AF009246"

/codon_start=1

/product="activator of G protein signaling"

/protein_id="AAD34206.1"

/db_xref="GI:4959038"

/translation="MKLAAMIKKMCPSDSELSIPAKNYRMVILGSKVGKTAIVSRF
 LTGFEDAYPTIEDHFKFYSGEYQILDITLDSGNHFFPAMRRILSLITVDVFIIV
 FSLNDRDSFBEVQLRQIILDTKSLKNTKENVDPVLCGNKGRDRDFRVVDQREI
 BOLVDDPQRCAYFEISAKNSLSDOMFRALFAMAKLPSEMSPLDHRKVSVOYCDVILH
 KKALRNKKLLRAGSGGGGPDGAFGIVAFARRPSVHSDLMYIREKASAGSQKDYK
 RCVIS"

BASE COUNT 422 a 546 c 483 g 289 t
 ORIGIN

Alignment Scores:

Pred. No.: 3.37e-127 Length: 1740
 Score: 1449.00 Matches: 281
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 99.93% Indels: 0
 DB: 9 Gaps: 0

US-09-709-103-1F1 (1-282) x AF069506 (1-1740)

Qy 1 MetLysLeuAlaAlaMetIleLysLysMetCysProSerAspSerGluLeuSerIlePro 20
 Db 146 ATGAAACTGGCGCGCATGATCAAGAAAGATGTGCCGAGCGACTCGAGTCTGAGTATCCG 205
 Qy 21 AlaLysAsnCysTyrArgMetValIleLeuGlySerSerLysValGlyThrAlaIle 40
 Db 206 GCCAAGAACTGATATCGCATGCTATCTCGCTCGTCCAAAGTGGCGAAGCGGCCATC 265
 Qy 41 ValSerArgPheLeuThrGlyArgPheGluAspAlaTyrThrProThrIleGluAspPhe 60
 Db 266 GTGTGCGGCTTCTCACCAGCGCGCTTCGAGGAGCGCTACACGCTTACCATCGAGACTTC 325
 Qy 61 HisArgLysPheTyrSerIleArgGlyGluValTyrGlnLeuAspIleLeuAspThrSer 80
 Db 326 CACCGCAAGTTCTACTCTCCATCGCGCGGAGGTCTACCGAGTCTGACATCTCTGACACGTCC 385
 Qy 81 GlyAsnHisProPheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIle 100
 Db 386 GGCAACACACCGCTTCCCGCCATCGCGCGCTCTCCATCTCTCACAGGAGACGTTTTCATC 445
 Qy 101 LeuValPheSerLeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGlnGln 120
 Db 446 CTGTGTGTTTCAGTCTGGACACCGCGACTCTCTTCGAGGAGGTGACGCGCTCAGGACGAG 505
 Qy 121 IleLeuAspThrLysSerCysLeuLysAsnLysThrLysGluAsnValAspValProLeu 140
 Db 506 ATCTCTGACACCAAGTCTTGCTCTCAAGAACAAACCAAGGAGAGACGTGGACGTCCCTG 565
 Qy 141 ValIleCysGlyAsnLysGlyAspArgAspPheTyrArgGluValAspGlnArgGluIle 160
 Db 566 GTCATCTCGGCAACAGGGGTGACCGGACTTCTACCGGAGGTGGACCGCGGAGATC 625
 Qy 161 GluGlnLeuValGlyAspAspProGlnArgCysAlaTyrPheGluIleSerAlaLysLys 180
 Db 626 GAGCAGCTGTGTGGCGGAGACCCCGCGCTGCGCTACTTCGAGATCTCGGCCAAGAAG 685

QY 181 AanserSerleuaspGlnmetPheargalaleuphealaleuAlaLysleuProserGlu 200
 Db 686 AACAGAGAGCTGACCAAGATGTTCCGCGCTCTTCCCATGCGCAAGCTGCCACGAG 745
 QY 201 MetSerProaspLeuHisArgLysValSerValGlnTyrCysaspValleuHisLysLys 220
 Db 746 ATGAGCCCAAGCTGACCAAGGCTCTGCTGAGTACTGACGCTGCTGACCAAGAG 805
 QY 221 AlaleuArgAsnLysLysLeuLeuArgAlaGlySerGlyGlyGlyGlyGlyGlyGly 240
 Db 806 GCGCTGCGAACAAGAGCTGCTGCGCGCGCGCAAGGCGGCGGCGGCGCAACCGGCG 865
 QY 241 AspalAphneGlyTlleValAlaProPheAlaArgProSerValHisSeraspLeuMet 260
 Db 866 GACGCTTTGGCATCGTGGACCTTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 925
 QY 261 TyrleArgGluLysAlaSerAlaGlySerGlnAlaLysAspGluArgCysValle 280
 Db 926 TACATCCCGAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTG 985
 QY 281 Ser 281
 Db 986 AGC 988

RESULT 5
 LOCUS BC018041 1758 bp mRNA linear PRI 06-DEC-2001
 DEFINITION Homo sapiens, RAS, dexamethasone-induced 1, clone MGC:26290
 IMAGE:4797482, mRNA, complete cds.

ACCESSION BC018041
 VERSION BC018041.1 GI:17390075
 KEYWORDS MGC.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 1758)
 STRAUSBERG, R.
 DIRECT SUBMISSION
 SUBMITTED (03-DEC-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA

REMARK
 COMMENT NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 CDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki
 Toshiyuki and Piero Carninci (RIKEN)
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome
 Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www.shgc.stanford.edu>
 Contact: (Dickson, Mark) mdedpaxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

FEATURES

source

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAX Plate: 32 Row: f Column: 11.
 Location/Qualifiers

1..1758
 /organism="Homo sapiens"
 /db_xref="locusID:51655"
 /db_xref="taxon:9606"
 /clone="MGC:26290 IMAGE:4797482"
 /tissue_type="Brain, hypothalamus"
 /clone_id="NIH_MGC_96"
 /lab_host="DH10B"
 /note="Vector: pBluescript"
 213..1058
 /codon_start=1
 /product="RAS, dexamethasone-induced 1"

/protein_id="AAH18041.1"
 /db_xref="GI:17390076"
 /translation="MKLAAMIKMCPSEBELSIPAKNCYRMVILGSSKVGKTAISRFL
 LGRPEDATPTIPIDPFRKFSIRGEVQOLDITDLSGHPFAMRRLSILIGDVFILV
 FSLDNDRSFEVQRLRQQLDPLKSLKNTKENVVPLVICGNKDRPFREVDREI
 EDLVGDPPORCAVFEISAKNSSLDPOMFALPALMKLPSBMSPDHRRVSVQYCVLH
 KRLNRKLLRLRAGSGGGGDDPDAGFIVAIFARPSVHSDIMYIREKASGQADKXE
 RCVIS"
 BASE COUNT 387 a 576 c 501 g 294 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 3,42e-127 Length: 1758
 Score: 1449.00 Matches: 281
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 99.93% Indels: 0
 DB: 9 Gaps: 0

US-09-709-103-1F1 (1-282) x BC018041 (1-1758)

QY 1 MetLysleuAlaMetLleLysLysMetCysProSeraspSerGluLeuSerllePro 20
 Db 213 ATGAAGCTGGCCGCGAGATCAAGAGATGTGCCAGAGACTCGAGCTGAGTATCCG 272
 QY 21 AlalysasnCysTyrArgMetVallleLysGlySerSerLysValGlySerThrAlaile 40
 Db 273 GCCAAGACTGCTNTGCCAGTGCATCTCTGCTCTCAAGTGGGCAAGACGCGCATC 332
 QY 41 ValSerArgPheLeuThrGlyArgPheGluAspAlaTyrThrProThrllleGluasphe 60
 Db 333 GTGTGCGGCTTCTCTCACCGGCGCGCTTCAGAGAGCGCTTACACGCTCATCGAGACTTC 392
 QY 61 HisArgLysPheTyrSerlleArgGlyGluValTyrGluLeuAspLleLeuSPTTTSer 80
 Db 393 CACCGAGTTCATCTCATCCGCGGCGAGGTCTACAGCTGCACATCTTCACACAGTCC 452
 QY 81 GlyAsnHisProPheProAlaMetArgArgLeuSerlleLeuThrGlyAspValPheile 100
 Db 453 GGCAACACCGCTTCCCGCGCAAGCGCGCTTCTCATCTTCACAGAGACGTTTCATC 512
 QY 101 LeuValPheSerLeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGln 120
 Db 513 CTGGTGTTCAGTCTGAGCAACCGGACTCCTTCAGAGAGGTGACGCGCTCAGGCGAG 572
 QY 121 lleLeuAspThrLysSerCysLeuLysAsnLysThrLysGluAsnValAspValPhele 140
 Db 573 ATCTTCGACACCAAGCTTGTGCTCAAGAAACAAGAGAGACGTGACGCTGCCCCCTG 632
 QY 141 VallleCysGlyAsnLysGlyAspArgAspPheTyrArgGluValAspGlnArgGluile 160
 Db 633 GTCATCTGGCGCAACAAGGTGACCGGACTTTCACCGGAGGTGACCAAGCGCGAGATC 692
 QY 161 GluGluLeuValGlyAspAspProGlnArgCysAlaTyrPheGluLleSerAlaLysLys 180
 Db 693 GACGAGCTGTGGGCGACGACCCCGAGCGCTGCGCTTCTGCAAGATCTCGGCAAGAG 752
 QY 181 AanserSerleuaspGlnmetPheargalaleuphealaleuAlaLysleuProserGlu 200
 Db 753 AACAGAGAGCTGACCAAGATGTTCCGCGCTCTTCCCATGCGCAAGCTGCCACGAG 812
 QY 201 MetSerProaspLeuHisArgLysValSerValGlnTyrCysaspValleuHisLysLys 220
 Db 813 ATGAGCCCAAGCTGACCAAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 872
 QY 221 AlaleuArgAsnLysLysLeuLeuArgAlaGlySerGlyGlyGlyGlyGlyGlyGly 240
 Db 873 GCGCTGCGAACAAGAGCTGCTGCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 932
 QY 241 AspalAphneGlyTlleValAlaProPheAlaArgProSerValHisSeraspLeuMet 260
 Db 933 GACGCTTTGGCATCGTGGACCTTCCGCGCGCGCGCGCGCGCGCGCGCGCGCTG 992

Qy 261 TyrIleArgGluLysAlaSerAlaGlySerGlnAlaLysAspLysGluArgCysValIle 280
 Db 993 TATATCCGCGAGAGGCGAGCGCGGAGCAAGCAAGGAGCGCTCGCTCATC 1052

Qy 281 Ser 281
 Db 1053 AGC 1055

RESULT 6
 AF177335 1746 bp mRNA linear PRI 03-OCT-2000
 LOCUS Homo sapiens clone SP1942 unknown mRNA.
 DEFINITION AF177335
 ACCESSION AF177335
 VERSION AF177335.1 GI:10503968
 KEYWORDS
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1746)
 AUTHORS Gu, J.R., Wan, D.F., Zhao, X.T., Zhou, X.M., Jiang, H.Q., Zhang, P.P.,
 Qin, W.X., Huang, Y., Qiu, X.K., Qian, L.F., He, L.P., Li, H.N., Yu, Y.,
 Yu, J. and Han, L.H.
 TITLE Novel human cDNA clone with function of inhibiting cancer cell
 growth
 JOURNAL
 REFERENCE 2 (bases 1 to 1746)
 AUTHORS Gu, J.R., Wan, D.F., Zhao, X.T., Zhou, X.M., Jiang, H.Q., Zhang, P.P.,
 Qin, W.X., Huang, Y., Qiu, X.K., Qian, L.F., He, L.P., Li, H.N., Yu, Y.,
 Yu, J. and Han, L.H.
 TITLE Direct Submission
 JOURNAL
 REFERENCE Submitted (06-AUG-1999) National Laboratory For Oncogenes & Related
 Genes, Shanghai Cancer Institute, 25/Ln 2200 Xie Tu Road, Shanghai
 200032, P.R. China
 FEATURES
 source
 1. 1746
 /organism="Homo sapiens"
 /db xref="taxon:9606"
 /clone="SP1942"
 455..1039
 /codon_start=1
 /product="unknown"
 /protein_id="AAG17979.1"
 /db xref="GI:10503968"
 /translations="MRLSLITGDVFLVSLNDRDSFEVORLRQIILDKSLKNNK
 TKNVDVPLVICNGKGRDREYVLDQDPPORCAVFEISAKNSSLQDMFR
 ALFMAKLPMSESPDLHRKVSQYCDVHLKALRNKLLRAGSGGGGDPGAFGIVA
 PFARRPSVHSDLMYIREKASAGQAKDKRCVLS"
 BASE COUNT 388 a 569 c 496 g 293 t
 ORIGIN

Alignment Scores:
 Pred. NO.: 5.7e-126 Length: 1746
 Score: 1436.00 Matches: 281
 Percent Similarity: 99.65% Conservative: 0
 Best Local Similarity: 99.65% Mismatches: 0
 Query Match: 99.03% Indels: 1
 DB: 0 Gaps: 0

US-09-709-103-1F1 (1-282) x AF177335 (1-1746)

Qy 1 MetLysLeuAlaAlaMetIleLysLysMetCysProSerAspSerGluLeuSerIlePro 20
 Db 193 ATGAACATGGCGCGATGATCAAGAGATGTGCCGAGCGACTCGGAGCTGAGTATCCCG 252

Qy 21 AlaLysAsnCysTyrArgMetValIleLeuGlySerSerLysValGlyLysThrAlaIle 40
 Db 253 GCCAAGAACTGCTATCGCATCGCTATCTCGGCTCGTCAAGGTGGGCAAGAGCGGCATC 312

Qy 41 ValSerArgPheLeuThrGlyArgPheGluAspAlaTyrThrProThrIleGluAspPhe 60
 Db 313 GTGTGCGCGCTTCCTCACCGCGCGCTTCGAGAGCGCTACACGCCTACATCGAGGACTTC 372

Qy 61 HisArgLysPheTyrSerIleArgGlyValTyrGlnLeu-AspIleLeuAspThrSe 80
 Db 373 CACGCAAGTTCTACTCATCCGCGGAGGTCTACCAGCTTCGACATCTCTCGACAGTC 432

Qy 80 rGlyAsnHisProPheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIle 100
 Db 433 CGGCNACCAACCCGTTCCCGCCATGCGGCGCTCTCCATCTCACAGAGAGCGTTTCAT 492

Qy 100 eLeuValPheSerLeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGlnIle 120
 Db 493 CTTGGTGTTCAGTCTGGACAACCGCGACTCTCTTCAGAGAGGTGCGCGGTCTCAGCGACA 552

Qy 120 nIleLeuAspThrLysSerCysLeuLysAsnLysThrLysGluAsnValAspValProLe 140
 Db 553 GATCCTCGACACCAAGTCTTCCTCAAGAACAAACCAAGAGAACCTGGAGCGTGCCTC 612

Qy 140 uValIleCysGlyAsnLysGlyAspArgAspPheTyrArgGluValAspGlnArgGluIle 160
 Db 613 GGTCTATCTGCGCAACAAGGTGACCGGACTTCTACCGGAGGTGGACCGCGAGAT 672

Qy 160 eGluGlnLeuValGlyAspAspProGlnArgCysAlaTyrPheGluLysSerAlaLysLy 180
 Db 673 CGAGCAGCTGTGGCGGACGACCCCGAGCGCTGCGCTACTTTCAGATCTCGGCAAGAA 732

Qy 180 sAsnSerSerLeuAspGlnMetPheArgAlaLeuPheAlaMetAlaLysLeuProSerGln 200
 Db 733 GAACAGCAGCTGGACCAAGATGTTCCGCGCTCTTCCCATGCGCAAGTGGCCAGCGA 792

Qy 200 uMetSerProAspLeuHisArgLysValSerValGlnTyrCysAspValLeuHisLysLy 220
 Db 793 GATGAGCCAGACCTGACCCCAAGGTCTCGGTGAGTACTGCGAGCTGCTGCAAGAA 852

Qy 220 sAlaLeuArgAsnLysLysLeuLeuArgAlaGlySerGlyGlyGlyGlyAspProGln 240
 Db 853 GCGCTGCGGAAACAAGAGCTGTGCGGCGCGCAGCGCGCGCGCGCGCGCGCGCGG 912

Qy 240 yAspAlaPheGlyLleValAlaProPheAlaArgArgProSerValHisSerAspLeuMe 260
 Db 913 CGACGCTTTGGCATCTGTCGACACCTTTTCGCGCGCGCGCGCGCGCGCGCGCGCTCAT 972

Qy 260 tTyrIleArgGluLysAlaSerAlaGlySerGlnAlaLysAspLysGluArgCysValIle 280
 Db 973 GTACATCCGAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGCTGCTCAT 1032

Qy 280 eSer 281
 Db 1033 CAGC 1036

RESULT 7
 BC034166 1612 bp mRNA linear ROD 07-AUG-2002
 LOCUS BC034166
 DEFINITION BC034166
 IMAGE:4989312, mRNA, complete cds.
 ACCESSION BC034166
 VERSION BC034166.1 GI:21706874
 KEYWORDS MGC.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 1612)
 AUTHORS Strausberg, R.
 TITLE Direct Submission
 JOURNAL
 REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 COMMENT Contact: MGC help desk
 Email: cgapbs-remail.nih.gov
 Tissue procurement: Jeffrey E. Green, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www-shgc.stanford.edu>
 Contact: (Dickson, Mark) mcdpax11.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILMUT at: <http://image.llnl.gov>
 Series: IRAP Plate: 58 Row: 9 Column: 5
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 6677672.

FEATURES

source

1. 1612
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /map="FVB/N"
 /clone="MGC:36188 IMAGE:4989312"
 /tissue_type="Colon, normal, 5 month old male mouse."
 /clone_11b="NCI CGAP_C024"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"
 108. 950
 /codon_start=1
 /product="RAS, dexamethasone-induced 1"
 /protein_id="AAH34166.1"
 /db_xref="GI:21706875"
 /translation="MKLAAMIKKMCPSDSELSIPAKNCYRWVILGSSKVKTAIVSR
 LTRGFEDATPTIEDFHRKFTSINGEYQDLITSGNHPNRRISILTGDFITV
 FSLVDRDSFEEVQRLKQIIDLTKSCLNKTKLENDVPLVTCNGKGDHPREVEQREI
 EQLVDDPQRCAYFEISAKNSLDQWRALEFMAKLPSEMSPDLHRKVSQVCDVH
 KKALRNKKILRAAGSGGGGHDGDAFGIILAFARPSVSHDIYIREKTSVSGQADKKR
 CIVIS"

CDS

BASE COUNT 417 a 424 c 420 g 351 t

ORIGIN

Alignment Scores:

Pred. No.: 3.88e-123 Length: 1612
 Score: 1405.50 Matches: 274
 Percent Similarity: 98.56% Conservative: 3
 Best Local Similarity: 97.51% Mismatches: 3
 Query Match: 96.93% Indels: 1
 DB: 10 Gaps: 1

US-09-709-103-1f1 (1-282) x BC034166 (1-1612)

QY 1 MetlySLuAlAlAlMetllySLySMetCySProSerASPSeTGlUleuSerllepPro 20
 Db 108 ATGAAACTGCGCGCATGATCAAGAAAGTGTCCCAAGGACTCTGAATGATCCCG 167
 QY 21 AlAlYsAnCySTyArGMeTVallleuGlYSerSeTlYsValGlylYThrAlle 40
 Db 168 GCCAAGAACTGCTACGATGATGATCTCTCGGCTCATCCAAAGTGCGCAAGCGCATC 227
 QY 41 ValSerArGpHeuThrGlyArGpHeGluAsePalAlyThrProThrlleGluAspHe 60
 Db 228 GTGTGGCGGTTCTCACGGCGGCTTCCAGATGCTTACACCCCTACATGAGGACTTC 287
 QY 61 HlSaIglYSPHeTYSerlLeArGlyGluVallyTgInLeuAspIlleuAspThlSer 80
 Db 288 CACCAAAAGTTTACTCGATCCGCGCGCAAGCTACACAGTGGACATCTGACACATCC 347
 QY 81 GlYAsnHlSPHeProHlaMeTArGArGleuSerlleuThrGlyAspValPheIlle 100
 Db 348 GGCATTCATCCGTTCCGCCCATGCGGCGCTCTCTATCTTACAGAGAGCGATTTCATT 407
 QY 101 LeuValPheSerleuAspAnArGAspSerPheGluGluValGlnArGleuArGlnGln 120
 Db 408 CTGGTGTTCAAGCTTAGCAACCGCATCTATTGAAAGAGGTCAAAAGCTCAACACAGCAG 467
 QY 121 lleleuAspThlYSerCySLeulYsAnlyThrlYgluAsnValAspValProleu 140

Db 468 ATCTTAGACACCAAGTCTGTCTCAAGAACAAACCAAGAGATGTGACGTGCGGCTG 527
 QY 141 VallleCySLyAnlySGlyAspArGAspPheTyrArGlyuValAspGlnArGlyuile 160
 Db 528 GTCAATTGCGGTTAAACAAGGGGACCGGACTTCTTACCGGAAGTAAAGCAGCGGAGATT 587
 QY 161 GluGlnleuValGlyAspAspProGlnArGlyAspAlaTyrPheGluileSerAlaYls 180
 Db 588 GAGCAGCTGGTGGGTGACGACCTCAGCGGTTGGCTTACATTCAGATCTCAGCAAGAAG 647
 QY 181 AsnSerSerleuAspGlnMetPheArGAlaIleuPheAlaMetLAlYsleuPProSerGlu 200
 Db 648 AACGACGCTTGACCAAGATTCCTGCGCTTTGCCATGCGCAAGCTGCTTACCGAG 707
 QY 201 MetSerProAspLeuHlSaIArglyValSerValGlnTyrCyAspValleuHlSaYls 220
 Db 708 ATGAGCCCACTTGACCGCAAGGATCTGTCTCAAGTACTGCACTGCAACAAAG 767
 QY 221 AlAluArGAnlySLySLeulYsleuArGAlaGlySerGlyGlyGlyAspProGly 240
 Db 768 GCTCTGAGAAACAAGAGCTTCTGCGGCGGCGAGC--GAGGCGGGGGCGACACGCG 824
 QY 241 AspAlaPheGlylleValAlaPProPheAlaArGArGProSerValHlSerAspLeuMet 260
 Db 825 GATGCTTTGGCATCTTGGCGCTTTGCTGCGCAACCAAGCTGCAACGCACTCATG 884
 QY 261 TyrilleArGlySLySLaSerAlaGlySerGlnAlaYsAspLySGluArGlyValille 280
 Db 885 TACATTGTAAAAAACCACTGTGCGGACGAGCTTAAGCAAGAGCGCTGTGCATC 944
 QY 281 Ser 281
 Db 945 AGT 947

RESULT 8

AF009246

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

AF009246 1623 bp mRNA ROD 13-FEB-1998
 AF009246 Mus musculus ras-related protein (DEXRASI) mRNA, complete cds.
 AF009246.1 GI:2253712
 Mus musculus.
 Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 Kempainen, R.U. and Behrend, E.N.
 Dexamethasone rapidly induces a novel ras superfamily
 member-related gene in AtT-20 cells
 Submitted (18-JUN-1997) Physiology & Pharmacology, Auburn
 University College of Veterinary Medicine, 213 Greene Hall, Auburn,
 AL 36849, USA

FEATURES

source

1. 1623

/organism="Mus musculus"

/db_xref="taxon:10090"

/cell_line="corticotrope tumor cell line AtT-20"

1. 1623

/gene="DEXRASI"

142. 984

/note="induced by dexamethasone"

/codon_start=1

/product="ras-related protein"

/protein_id="AAC53538.1"

/db_xref="GI:2253713"

/translation="MKLAAMIKKMCPSDSELSIPAKNCYRWVILGSSKVKTAIVSR"

LTGRFEDAYTPTIEDFHRKFYSIRGEVYQDILDTSNGNHPFPAMRRLSILTGDFVILV
 FSLNDRDSFEVORLQKQILDTSKLNKTKENVDPLVICNGKGDYDFREVEQREI
 EQLVGGDDPQRCAYFEISAKNSSLQDMFRALFAMAKLPSEMSPLDHRKVSQYCDVLH
 KKAALRNKLLRAGSGGGGDHDAFGLAPFARRPSVHSDLMYIREKTSVGSQAKDKER
 CVIS"

BASE COUNT 412 a 437 c 417 g 357 t
 ORIGIN

Alignment Scores:

Pred. No.: 3,91e-123 Length: 1623
 Score: 1405.50 Matches: 274
 Percent Similarity: 98.58% Conservative: 3
 Best Local Similarity: 97.51% Mismatches: 3
 Query Match: 96.93% Indels: 1
 DB: 10 Gaps: 1

US-09-709-103-1F1 (1-282) x AF009246 (1-1623)

Qy 1 MetLysLeuAlaMetIleLysLysMetCysProSerAspSerGluLeuSerIlePro 20
 Db 142 ATGAACCTGGCGCGATGATCAAGAAGATGTGCCAAGCGACTCTGAACCTGATCCCG 201
 Qy 21 AlalysAsnCysTyrArgMetValIleLeuGlySerSerLysValGlyLysThrAlaIle 40
 Db 202 GCCAAGAAGCTCTACAGGATGTGATCTCGGCTCATCCAAAGTGGCAAGCGCCATT 261
 Qy 41 ValSerArgPheLeuThrGlyArgPheGluAspAlaTyrThrProThrIleGluAspPhe 60
 Db 262 GTGTCGCGCTTCTCAGCGCGGCTTTCAGGATGCTTACACCTTACATCGAGGACTTC 321
 Qy 61 HisArgLysPheTyrSerIleArgGlyGluValTyrGlnLeuAspIleLeuAspThrSer 80
 Db 322 CACCGAAGTTTACTCGATCCGCGCGAAGTCTACAGTTGGACATCTGGACACATCC 381
 Qy 81 GlyAsnHisProPheProAlaMetArgArgLeuSerIleLeuThrGlyAspValPheIle 100
 Db 382 GGCAATCATCGCTTCCCGCATCGCGCGCTCTCTATCTCTCACAGAGACGTTTTCATT 441
 Qy 101 LeuValPheSerLeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGlnGln 120
 Db 442 CTGGTGTTCAGCTTAGACAACCGGACTCATTCGAAGAGTGCAGAGGCTCAACAGCAG 501
 Qy 121 IleLeuAspThrLysSerCysLeuLysAsnLysThrLysGluAsnValAspValProLeu 140
 Db 502 ATCTAGACACCAAGTCTGTCTCAAGAACAAACCAAGAGATGTGGAGTGCCTG 561
 Qy 141 VallIleCysGlyAsnLysGlyAspArgAspPheTyrArgGluValAspGlnArgGluIle 160
 Db 562 GTCATTTTCGGTAAACAAAGGGGACCGGACTTCTACCGGGAAGTAGACGCGGAGATT 621
 Qy 161 GluGlnLeuValGlyAspAspProGlnArgCysAlaTyrPheGluIleSerAlaLysLys 180
 Db 622 GAGCAGCTGGTGGTGAGACCTCAGCGTTGTGCTACTTCGAGATCTCAGCCAGAG 681
 Qy 181 AsnSerSerLeuAspGlnMetPheArgAlaLeuPheAlaMetAlaLysLeuProSerGlu 200
 Db 682 AACAGCAGCTTGGACACAGATGTTCGTCGCTCTTTGCGCATGGCCAAAGCTGCTAGCGAG 741
 Qy 201 MetSerProAspLeuHisArgLysValSerValGlnTyrCysAspValLeuHisLysLys 220
 Db 742 ATGACCCCGAGCTTGCACCGAAGTATCTGTGCAGTACTGCGACGTAAGTGCACAGAG 801
 Qy 221 AlaLeuArgAsnLysLysLeuLeuArgAlaGlySerGlyGlyGlyGlyAspProGly 240
 Db 802 GCTCTGAGGALACAGAGACTTCTGCTGCGGGCAGC---GGAGCGGGGGCGACCGCC 858
 Qy 241 AspAlaPheGlyIleValAlaProPheAlaArgProSerValHisSerAspLeuMet 260
 Db 859 GATGCTTTGGCATCTTGGCCCTTTGCTCGCAGACCCAGCGTGCACAGCGACTCATG 918
 Qy 261 TyrIleArgGluLysAlaSerAlaGlySerGlnAlaLysAspLysGluArgCysValIle 280
 Db 919 TACATTCGTGAAAAAACACAGTGTCTGGCAGCAGGCTAAGGACAAAGGACGCTGTGTCATC 978

Qy 281 Ser 281
 Db 979 ACT 981

RESULT 9

AF239157 1616 bp mRNA linear ROD 06-JUL-2001
 LOCUS Rattus norvegicus DEXRAS1 (Dexras1) mRNA, complete cds.

DEFINITION AF239157

ACCESSION AF239157

VERSION AF239157.1 GI:7230767

KEYWORDS

SOURCE

ORGANISM

Rattus norvegicus.

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 1616)

Fang, M., Jaffrey, S.R., Sawa, A., Ye, K., Luo, X. and Snyder, S.H.

Dexras1: a G protein specifically coupled to neuronal nitric oxide

synthase via CAPON

Neuron 28 (1), 183-193 (2000)

20537828

11086993

2 (bases 1 to 1616)

Fang, M., Jaffrey, S.R., Sawa, A., Ye, K. and Snyder, S.H.

Direct Submission

Submitted (24-FEB-2000) Neuroscience, Johns Hopkins University, 725

N. Wolfe St, Baltimore, MD 21205, USA

Location/Qualifiers

1..1616

/organism="Rattus norvegicus"

/db_xref="taxon:10116"

1..1616

/gene="Dexras1"

/gene="Dexras1"

/note="Ras-related protein; small GTPase; similar to Mus

musculus DEXRAS1 encoded by GenBank Accession Number

AF009246"

/codon_start=1

/product="DEXRAS1"

/protein_id="AAF43090.1"

/db_xref="GI:7230768"

/translation="MKLAAMTKMCPDSELSIPAKNCYRMVILGSSKVGKTAIVSRF

LTGRFEDAYTPTIEDFHRKFYSIRGEVYQDILDTSNGNHPFPAMRRLSILTGDFVILV

FSLNDRDSFEVORLQKQILDTSKLNKTKENVDPLVICNGKGDYDFREVEQREI

EQLVGGDDPQRCAYFEISAKNSSLQDMFRALFAMAKLPSEMSPLDHRKVSQYCDVLH

KKAALRNKLLRAGSGGGGDHDAFGLAPFARRPSVHSDLMYIREKTSVSSQAKDKER

CVIS"

BASE COUNT 403 a 436 c 413 g 363 t 1 others

ORIGIN

Alignment Scores:

Pred. No.: 1.43e-122 Length: 1616

Score: 1399.50 Matches: 273

Percent Similarity: 98.22% Conservative: 3

Best Local Similarity: 97.15% Mismatches: 4

Query Match: 96.52% Indels: 1

DB: 10 Gaps: 1

US-09-709-103-1F1 (1-282) x AF239157 (1-1616)

Qy 1 MetLysLeuAlaMetIleLysLysMetCysProSerAspSerGluLeuSerIlePro 20

Db 118 ATGAACCTGGCGCGATGATCAAGAAGATGTGCCAAGCGACTCTGAACCTGATCCCG 177

Qy 21 AlalysAsnCysTyrArgMetValIleLeuGlySerSerLysValGlyLysThrAlaIle 40

Db 178 GCCAAGAAGCTCTACAGGATGTGATCTCGGCTCATCCAAAGTGGCAAGCGCCATC 237

Qy 41 ValSerArgPheLeuThrGlyArgPheGluAspAlaTyrThrProThrIleGluAspPhe 60

```

Db 238 GTGTCGCGCTTCTCAACGGCGCGCTTCGAGAGACGCTTACACCCCTTACCATTTGAGACCTTC 297
Qy 61 H1sarglysePheTyrSerTleArglygluValTyrGlnLeuAspTleuAspThrSer 80
Db 298 CACCGAAAGTTTACTCGATCCGCGGCAAGTCTACAGTACGACATGACATGACACATCT 357
Qy 81 G1AsnHisProPheProAlaMetArgArgLeuSerTleLeuThrGlyAspValPheIle 100
Db 358 GGCATTCATCCGTTTCCCGCCATGGCGGCGCTCTCATCTCACAGGAGACGTTTCACTT 417
Qy 101 LeuValPheSerLeuAspAsnArgAspSerPheGluGluValGlnArgLeuArgGlnGln 120
Db 418 CTGCTGTTGAGCTTACCAACCGCGACTCTTCAAGAGAGCTCAAGGCTCAACACAGCAG 477
Qy 121 TleLeuAspThrLysSerCysLeuLysAsnLysThrLysGluAsnValAspValProLeu 140
Db 478 ATCTTAGACACCAAGTCTGCTCAAGAACAAACCAAGAGAGATGAGACGTCGCGCTG 537
Qy 141 ValIleCysGlyAsnLysGlyAspArgAspPheTyrArgGluValAspGlnArgGluIle 160
Db 538 GTCATTTGGGTACCAAGGAGGACCGGACCTTACCGGAGATGAGAGCGGAGATT 597
Qy 161 G1UGlnLeuValGlyAspAspProGlnArgCysAlaTyrPheGluIleSerAlaLysLys 180
Db 598 GAGCACTGCTGGCGCATGACCTTCAAGCTTGTGCTTCAAGATCTCGGCGCAAGAG 657
Qy 181 AsnSerSerLeuAspGlnMetPheArgAlaLeuPheAlaMetAlaLysLeuProSerGlu 200
Db 658 AATAGCAGCTTGACCAAGATGTTCCGTCGCTTTCCTTCCATGAGCCCAAGGCTGACGAG 717
Qy 201 MetSerProAspLeuHisArgLysValSerValGlnTyrCysAspValLeuHisLysLys 220
Db 718 ATGAGCGCTGACCTTGACCGCAAGGTGTCTGTGCACTGACGTGACGACAAAAG 777
Qy 221 AlaLeuArgAsnLysLysLeuLeuArgAlaGlySerGlyGlyGlyGlyGlyAspProGly 240
Db 778 GCTCTAGAGAACAAAGAGCTTCTGCTGCGGCGAGC---GAGAGTGGGCGCACCAAG 834
Qy 241 AspAlaPheGlyLysValAlaProPheAlaArgArgProSerValHisSerAspLeuMet 260
Db 835 GATGCTTTGGCATCTTGGCGCTTGTCTGCGAGACTGACGATGACGACCTCATG 894
Qy 261 TyrIleArgGlyLysAlaSerAlaGlySerGlnAlaLysAspLysGluArgCysValIle 280
Db 895 TACATTCGTGAGAAACCAAGTCTCAGACAGCAGGCTTAAGAGCAAGAGCGCTGTGCATC 954
Qy 281 Ser 281
Db 955 AGT 957

RESULT 10
AF222979 4990 bp DNA linear PRI 02-JAN-2001
LOCUS Homo sapiens activator of G-protein signaling gene, complete cds.
DEFINITION AF222979
ACCESSION AF222979
VERSION AF222979.1 GI:12004991
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE
1 (bases 1 to 4990)
AUTHORS Cismowski,M.J., Xie,X. and Duzic,E.
TITLE Genomic sequence of the human ras-related G-protein activator AGS1
JOURNAL Unpublished
2 (bases 1 to 4990)
AUTHORS Cismowski,M.J., Xie,X. and Duzic,E.
TITLE Direct Submission
JOURNAL Submitted (10-JAN-2000) OSI Pharmaceuticals, 777 Old Saw Mill River
Road, Terrytown, NY 10591, USA
FEATURES
Location/Qualifiers
1..4990
/organism="Homo sapiens"

```

```

/db xref="taxon:9606"
/chromosome="17"
/map="17p11.2"
/clone="BamHI-EcoRI subclone of P1"
/notes="chromosomal localization verified by fluorescence
in-situ hybridization analysis"
join(2762..3188,3400..4643)
/product="activator of G-protein signaling"
join(2903..3188,3400..3959)
/notes="ras-related protein; AGS1"
/codon_start=1
/product="activator of G-protein signaling"
/protein_id="AA044256.1"
/db xref="GI:12004992"
/translation="MKLAMIKKMCPDSELSIPANCYRYMTILSSKYGVTAIVSRF
LNGRPDVAITPTEIDPFRKFSIRGVYDLDLDSGNHPPAMRSLITGDFITLV
FSLDNRDSFEVQRLRQQLIDPKSLCKKATKKNVPLVITCKNGRDPREVDREI
EDLVGDPPRCAYFETISAKNSLDQMFETAPAMAKLSKMSPDJHRKVSVDYVLA
KALRNKKLLRAGSGGGGDDPDARFQIVAPFARBSVHSDLMYIREKASGQAKDE
RCVIS"
BASE COUNT 1049 a 1507 c 1509 g 925 t
ORIGIN
Alignment Scores:
Pred. No.: 8,38e-121 Length: 4990
Score: 1387.00 Matches: 281
Percent Similarity: 79.83% Conservative: 0
Best Local Similarity: 79.83% Mismatches: 71
Query Match: 95.66% Indels: 71
Gaps: 1
US-09-709-103-1f1 (1-282) x AF222979 (1-4990)
Qy 1 MetLysLeuAlaAlaMetIleLysLysMetCysProSerAspSerGluLeuSerTlePro 20
Db 2903 ATGAACTGGCGCGAGATGATCAAGAAATGTCGCCGAGCACTGGAGCTGAGTATCCG 2262
Qy 21 AlaLysAsnCysTyrArgMetValIleLeuGlySerSerLysValGlyLysThrAlaIle 40
Db 2963 GCCAAGAACTCTTATCCCATGTCATCTCGGCTCGTCCAAAGTGGCAAGAGCGGCATC 3022
Qy 41 ValSerArgPheLeuThrGlyArgPheGluAspAlaTyrThrProThrIleGluAspPhe 60
Db 3023 GTGTGCGCTTCTTCAACCGCGCTTCAAGAGCGCTTCAACAGCTTCAACATGAGACTTC 3082
Qy 61 HisArgLysPheTyrSerTleArgGlyGluValTyrGlnLeuAspTleLeuAspThrSer 80
Db 3083 CACCGCAAGTTCTACTCCATCCGCGGAGAGTTCACAGCTCGACATCTCGACAGCTTC 3142
Qy 81 G1AsnHisProPheProAlaMetArgArgLeuSerTleLeuThr----- 95
Db 3143 GGCACACACCGCTTCCCGCATCGCGGCTCTTCATCTTCACAGGTGAGCGGGGGCC 3202
Qy 95 ----- 95
Db 3203 GGGCAGTGGCGGAGGAGGAGGCGGGAACCTTCGCGCAGGCGCCCCGAGAGCGCGGT 3262
Qy 95 ----- 95
Db 3263 CCGGCTGCGCGCGCGCGAGTAGTGCGCTTGCCTTAGAGAGCTAGCGCGCGCGCG 3322
Qy 95 ----- 95
Db 3323 GCTCTCAAGTAGCGCCGACCTTGTCCCTGGCGGCGCACACCTTCTCTCTTCTGCT 3382
Qy 96 -----G1AspValPheIleLeuValPheSerLeuAspAsnArgAspSe 110
Db 3383 CTGTGTGCCCCCTTAGAGAGCGTTTTCATCTGTGTTCATGCTGAGACACCGCGCATC 3442
Qy 110 rPheGluGluValGlnArgLeuArgGlnGlnIleLeuAspThrLysSerCysLeuLys 130
Db 3443 CTTCGAGGAGGTGACGCGCTCAGCGACGAGATCTCGACACCAAGCTTTCCTTAAGAA 3502

```

```

QY 130 nlysthrlysgluasnvalaspvalproleuvalillecysglyasnlysglyaspargas 150
Db 3503 CAAACCAAGGAGACGTGGACGTGCCCTCGTTCATCTGCGGCAACCAAGGAGTACCAGCGA 3562
QY 150 pPheTyArgGluValAspGlnAtrGluileGluGlnLeuValGlyAspAspProGlnAr 170
Db 3563 CTTCTACCGGAGGTGGACCGGCGAGATCGAGCAGCTGGTGGCGGACGCCCGGCGG 3622
QY 170 gCysAlaTyRheGluileSerAlaLyLysAlaAsnSerSerLeuAspGlnMetPheArgAl 190
Db 3623 CTGCGCCTACTTCGAGATCTCGGCCAAGAACAGACAGCAGCCTGGACCATGATGTTCCGCGC 3682
QY 190 aleuPheAlaMetAlaLyLysLeuProSerGluMetSerProAspLeuHisArgLyLysValse 210
Db 3683 GCTCTTCGCCATGGCCCAAGCTGCCAGCGAGATGAGCCACAGCTGACCGCAAGGTCTC 3742
QY 210 rValGlnTyRyCysAspValLeuHisLyLysAlaLeuArgAsnLyLysLeuLeuArgAl 230
Db 3743 GGTGAGTACTGCGAGCTGCTGCACAGAGAGCGCTGCCGGAACAAGAGCTGCTGCCGCGC 3802
QY 230 aGlySerGlyGlyGlyGlyAspProGlyAspAlaPheGlyIleValAlaPropheAl 250
Db 3803 CGCAGCGCGCGCGCGCGCGAGCCCGCGCGAGCGCTTTGCGATCGTGGCACCCTTCGC 3862
QY 250 aArgArgProSerValHisSerAspLeuMetTyRleAArgGluLyLysAlaSerAlaGlySe 270
Db 3863 CGCGCGCGCGCGCGAGCTGACAGCAGCTCATGTACATCGCGAGAGCGCGCGCGCAG 3922
QY 270 rGlnAlaLyLysAspLyGluArgCysValIleSer 281
Db 3923 CCAGCGCCAGACAGAGAGCGCTGGTGCATCAGC 3956

RESULT 11
LOCUS AC020558 183334 bp DNA linear PRI 09-JAN-2002
DEFINITION Homo sapiens BAC clone RP11-524F11 from 17, complete sequence.
ACCESSION AC020558
VERSION AC020558.4 GI:13242397
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 183334)
AUTHORS Sulston, J.E. and Waterston, R.
TITLE Toward a complete human genome sequence
JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE 99063792
PUBMED 9847074
REFERENCE 2 (bases 1 to 183334)
AUTHORS Mulvaney, E., Maupin, R., LaPlant, Y. and Bielicki, L.
TITLE The sequence of Homo sapiens BAC clone RP11-524F11
JOURNAL Unpublished (2001)
REFERENCE 3 (bases 1 to 183334)
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (03-JAN-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 4 (bases 1 to 183334)
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (07-MAR-2001) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 5 (bases 1 to 183334)
AUTHORS Waterston, R.
TITLE Direct Submission
JOURNAL Submitted (09-JAN-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Mar 7, 2001 this sequence version replaced gi:7631041.
-----
Center: Washington University Genome Sequencing Center

```

Center code: WUGSC
 Web site: <http://genome.wustl.edu/gsc>
 Contact: sapiens@watson.wustl.edu
 ----- Summary Statistics -----
 Center project name: H_NH0524F11

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RP11-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P.Y., Zhao, B., Frengen, E., Taten, M., Catanese, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)

VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is AC073621. Actual start of this clone is at base position 1 of RP11-524F11; actual end is at base position 183334 of RP11-524F11.

The sequence from position 1987 to 2253 was derived from PCR product of RP11-524F11 BAC DNA.

FEATURES

	Location/Qualifiers	source
	1..183334	
	/organism="Homo sapiens"	
	/db_xref="taxon:9606"	
	/chromosome="17"	
	/map="17"	
	/clone="RP11-524F11"	
	/clone_lib="RPC1-11"	
misc_feature	1..319	
	/note="similar to EST AI672824 (NID:g4852555) we73e06.x1"	
misc_feature	339..887	
	/note="similar to EST AI968384 (NID:g5765202) wu02e05.x1"	
misc_feature	339..738	
	/note="similar to EST AI289781 (NID:g39331622) qw12a07.x1"	
misc_feature	339..658	
	/note="similar to EST AW072137 (NID:g6027135) wz98e07.x1"	
misc_feature	340..767	
	/note="similar to EST AI014398 (NID:g3228779) ou93b12.s1"	
misc_feature	340..686	
	/note="similar to EST AW237782 (NID:g6570171) xm81e08.x1"	
misc_feature	346..742	
	/note="similar to EST AI359064 (NID:g4110685) qy25e07.x1"	
misc_feature	405..884	
	/note="similar to EST AI279211 (NID:g3917445) qm24c09.x1"	
misc_feature	782..1335	
	/note="similar to EST BF042464 (NID:g10759519)"	
misc_feature	822..1319	
	/note="similar to EST BE864687 (NID:g10385980)"	
misc_feature	834..2931	
	/note="CpG_island (%C=69.0, o/e=0.80, #CpGs=226)"	


```

misc_feature 1135..1585 /note="similar to EST BF555582 (NID:g11665312)"
misc_feature 1168..1583 /note="similar to EST AM028127 (NID:g5886883) wv26c07.x1"
misc_feature 1432..1583 /note="similar to EST BE390509 (NID:g9335874)"
misc_feature 1440..1583 /note="similar to EST AW915326 (NID:g8081017)"
misc_feature 1450..1583 /note="similar to EST AA790463 (NID:g2850583) vw04g12.r1"
misc_feature 1795..2255 /note="similar to EST BE390509 (NID:g9335874)"
misc_feature 1795..2201 /note="similar to EST AA790463 (NID:g2850583) vw04g12.r1"
misc_feature 1795..2153 /note="similar to EST AW915326 (NID:g8081017)"
misc_feature 1795..2114 /note="similar to EST AM028127 (NID:g5886883) wv26c07.x1"
misc_feature 1804..2071 /note="similar to EST BE836084 (NID:g10268462)"
misc_feature 1817..2201 /note="similar to EST AA672515 (NID:g644732) v059e02.r1"
misc_feature 1844..2201 /note="similar to EST AM463433 (NID:g7033601)"
misc_feature 1985..2285 /note="similar to EST AA852097 (NID:g2940690)"
misc_feature 2538..3033 /note="similar to EST BF057610 (NID:g10811506)"
misc_feature 2554..2806 /note="similar to EST BF195653 (NID:g11082770)"
misc_feature 3561..3924 /note="similar to EST R18261 (NID:g711871) YG01a05.r1"
repeat_region 4027..4085 /rpt_family="MERL_type"
misc_feature 4670..5117 /note="similar to EST R42241 (NID:g817108) yf98b09.s1"
misc_feature 4857..5117 /note="similar to EST T16275 (NID:g518437)"
repeat_region 6155..6307 /rpt_family="Alu"
repeat_region 6309..6446 /rpt_family="L2"
repeat_region 6583..6891 /rpt_family="Alu"
repeat_region 6960..7153 /rpt_family="L1"
repeat_region 7525..7838 /rpt_family="Alu"
repeat_region 7851..8487 /rpt_family="L1"
repeat_region 9000..9104 /rpt_family="L2"
misc_feature 11463..11750 /note="similar to EST BF344873 (NID:g11292331)"
misc_feature 11465..11750 /note="similar to EST AI340185 (NID:g4077112) qo79d07.x1"
misc_feature 11465..11750 /note="similar to EST AI571390 (NID:g4534764) cn4se12.x1"
misc_feature 11465..11750 /note="similar to EST AI701138 (NID:g4989038) we10d09.x1"
misc_feature 11465..11750 /note="similar to EST AI803637 (NID:g5369109) tc43d01.x1"
misc_feature 11465..11750 /note="similar to EST BE747895 (NID:g10161887)"
misc_feature 11465..11750 /note="similar to EST BF203429 (NID:g11097015)"
misc_feature 11465..11697 /note="similar to EST AI869572 (NID:g5543540) w198d03.x1"
misc_feature 11467..11750 /note="similar to EST AI817466 (NID:g5436545) wk38g01.x1"
misc_feature 11467..11619 /note="similar to EST AA367144 (NID:g2019483)"
misc_feature 11468..11750 /note="similar to EST BE218333 (NID:g8905651) hv37c09.x1"
misc_feature 11468..11750 /note="similar to EST BE549542 (NID:g9791256)"
misc_feature 11468..11750 /note="similar to EST BE750272 (NID:g10164264)"
misc_feature 11478..11750 /note="similar to EST BE667874 (NID:g11941769)"
misc_feature 11480..11717 /note="similar to EST R34866 (NID:g791767) yh86f06.s1"
misc_feature 11483..11750 /note="similar to EST AA429062 (NID:g2110588) zv49a05.r1"
misc_feature 11541..11750 /note="similar to EST W38322 (NID:g1320126) zb19g06.r1"
misc_feature 11552..11750 /note="similar to EST BE910310 (NID:g10406772)"

Alignment Scores:
Pred. No.: 6.38e-119 Length: 183334
Score: 1387.00 Matches: 281
Percent Similarity: 79.83% Conservative: 0
Best Local Similarity: 79.83% Mismatches: 0
Query Match: 95.66% Indels: 71
DB: 9 Gaps: 1

US-09-709-103-1f1 (1-282) x AC020558 (1-183334)
QY 1 MetLysLeuAlaIleMetIleLysLysMetCysProSerAspSerGluLeuSerIlePro 20
DB 2080 ATGAACCTGGCGCGGATGATCAAGAAATGTGCCGACCACTCGAGCTGATATCCG 2021
QY 21 AlAlaLysAncCysTyrArgMetValIleLeuGlySerSerLysValGlyLysThrAlaIle 40
DB 2020 GCCAAGAACTGCTATCCAGATGATCTCGCTCGTCCCAAGGCGGCAAGCGGCATC 1961
QY 41 ValSerArgPheLeuThrArgPheGluAspAlaThrProThrIleGluAspPhe 60
DB 1960 GTGTGGCGCTTCTCCACCGCGCGCTTGAAGAGCCCTAACCGCTTACATCGAGACTTC 1901
QY 61 HisArgLysPheTyrSerIleArgGlyValTyrGlnLeuAspIleLeuAspThrSer 80
DB 1900 CACCGCAAGTCTTCTCATTCGCGCGGAGAGTCTTACAGCTTCACATCTCGACACGCTCC 1841
QY 81 GlyAsnHisProPheProAlaMetArgArgLeuSerIleLeuThr 95
DB 1840 GGCACACACCGCTTCCCGCATCGCGGCGCTCTCCATCTCACAGGTAGCGCGGAGCC 1781
QY 95 95
DB 1780 GGGCAGTGGCGGAGGAGGAGCGGGAACCTTCGCGCAGCGCGCCCGAGCGCGGT 1721
QY 95 95
DB 1720 CCGGCTGCGCGCGCGCGAGTATGCGCTTGGAGAGCTTACGCGCGCGCGCG 1661
QY 95 95
DB 1660 GCTCAAGTCAAGCCGACTTGTCCCTGGCGGCGCACCTTCACTTCTCTTTCTGTCT 1601
QY 96 -----GlyAspValPheIleLeuValPheSerLeuAspAsnArgAspSe 110
DB 1600 CTCTGTGCCCCCTTAGAGAGCGTTTATCTCGGTGTGAGTGTGACAAACCGGCACTC 1541
QY 110 PheGluGluValGlnArgLeuArgGlnGlnIleLeuAspThrLysSerCysLeuLysAs 130
DB 1540 CTTGAGAGAGGTGACGCGCTCAGGCAAGCATCTCGACACCAAGCTTGTGCTCAAGAA 1481
QY 130 PheThrLysGluAsnValAspValProLeuValIleCysGlyAsnLysGlyAspArgAs 150
DB 1480 CAAACCAAGAGAAACGTGAGCGTCCCTGTGATCTGCGGCAACAGAGTGAACCGCGA 1421
QY 150 PheThrArgGluValAspGlnArgGluIleGluGlnLeuValGlyAspAspProGlnArg 170
DB 1420 CTTTACCGGAGGTGACGCGCGAGATCGACACACTGTGTGGGCAACACCCCGCAGCG 1261

```

QY 170 gCyalaTyPheGluIleSerAlaLysLysAsnSerSerLeuAspGlnMetPheArgAl 190
 Db 1360 CTGCGCCTACTTCGAGATCTCGCGCAAGAAGAACAGCAGCTGGACACAGATGTTCCGCGC 1301
 QY 190 aLeuPheAlaMetAlaLysLeuProSerGluMetSerProAspLeuHisArgLysValSe 210
 Db 1300 GCTCTTGGCCATGGCCAGCTGCCAGCGAGATGAGCCAGACCTGCGACCCCAAGGTTCTC 1241
 QY 210 rValGlnTyrcysAspValLeuHisLysLysAlaLeuArgAsnLysLysLeuLeuArgAl 230
 Db 1240 GGTGAGTACTCGAGCTGCTGCACAAGAGAGCGCTGCGGAACAAGAGCTGCTGCGGC 1181
 QY 230 aclySerGlyGlyGlyGlyAspProGlyAspAlaPheGlylleValAlaProPheAl 250
 Db 1180 CGCGACGCGCGCGCGCGCGAGCCCGCGCGACCCCTTGGCATCGTGGCACCCCTTCGC 1121
 QY 250 aArgProSerValHisSerAspLeuMetTyrlleArgGluLysAlaSerAlaGlySe 270
 Db 1120 GCGCGCGCGCGCGAGCTGACAGAGCTCATGTACATCCGAGAAGGCGCGCGCGAG 1061
 QY 270 rGlnAlaLysAspLysGluArgCysValIleSer 281
 Db 1060 CCAGGCCAAGGACAAGGCGCTGGTCATCAGC 1027
 RESULT 12
 AF262018 5141 bp DNA linear PRI 31-MAY-2000
 LOCUS Homo sapiens dexamethasone-induced ras-related protein 1 (DEXRASI)
 DEFINITION gene, complete cds.
 ACCESSION AF262018
 VERSION AF262018.1 GI:8118456
 KEYWORDS
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 5141)
 AUTHORS Kempainen,R.J. and Behrend,E.N.
 JOURNAL Human Dextral gene
 REFERENCE 2 (bases 1 to 5141)
 AUTHORS Kempainen,R.J. and Behrend,E.N.
 TITLE Direct Submission
 JOURNAL Submitted (30-APR-2000) Anatomy & Physiology, Auburn University,
 College of Veterinary Medicine, Auburn, AL 36832, USA
 FEATURES
 source
 1. 5141
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="2"
 /map="2q32"
 /gene="DEXRASI"
 /gene="DEXRASI"
 join(<2548..2833,3045..>3604)
 /gene="DEXRASI"
 /product="dexamethasone-induced ras-related protein 1"
 join(2548..2833,3045..3604)
 /gene="DEXRASI"
 /note="Dextral"
 /codon_start=1
 /product="dexamethasone-induced ras-related protein 1"
 /protein_id="AAF72997.1"
 /db_xref="GI:8118457"
 /translation="MKLAAMIKKMCPDSELSIPAKNCYRMVILGSSKVKTAIVSRF
 LTGRFEDATPTTDFHRKFYSIRGEVYQDILDITSGNHPPFAMRRLSLITGVDVFLV
 FSLDNRDSEFVORLKOIIDLTKSLNKKTNVDVPLVICGNKGRDRFYRVDDOREI
 EQLNRDPQRCAYFEISAKNSLIDOMERAFAMKLPSEMSPLHRKYSVQCVDLH
 KKALRNKKLRASGGGDPGAFGVAFARRPSVHSDLMYIREKASAGSQAKDKE
 RCVIS"
 BASE COUNT 1090 a 1579 c 1522 g 950 t
 ORIGIN

Alignment Scores:
 Pred. No.: 1.67e-120 Length: 5141
 Score: 1384.00 Matches: 280
 Percent Similarity: 79.83% Conservative: 1
 Best Local Similarity: 79.55% Mismatches: 0
 Query Match: 95.45% Indels: 71
 DB: 9 Gaps: 1
 US-09-709-103-1f1 (1-282) x AF262018 (1-5141)
 QY 1 MetLysLeuAlaAlaMetIleLysLysMetCysProSerAspSerGluLeuSerIlePro 20
 Db 2548 ATGAATACTGGCCGCGATGATCAAGAAGATGTGCCGAGGAGCTCGAGCTGAGTATCCCG 2607
 QY 21 AlaLysAsnCysTyArgMetValIleLeuGlySerSerLysValGlyLysThrAlaIle 40
 Db 2608 GCCAAGAACTGCTATCGCATGTGTCCTCGCTCGTCCAAGTGGCAAGACGCCATC 2667
 QY 41 ValSerArgPheLeuThrGlyArgPheGluAspAlaTyThrProThrIleGluAspPhe 60
 Db 2668 GTGTCCGCTTCTCACCAGCGCTTCGAGGAGCGCTACACGCTTACCATCGAGACTTC 2727
 QY 61 HisArgLysPheTySerIleArgGlyGluValTyGlnLeuAspIleLeuAspThrSer 80
 Db 2728 CACCGCAAGTTCTATCCCATCCGCGGAGGTCTACAGCTCGACATCTCTGACACGTC 2787
 QY 81 GlyAsnHisProPheProAlaMetArgArgLeuSerIleLeuThr----- 95
 Db 2788 GGCAACACCCCGTTCCCGCCCATGCGGCGCTCTCCATCTCACAGGTGAGCGCGGGGCC 2847
 QY 95 ----- 95
 Db 2848 GGGCAGGTCCGGAGGGAAGGGCGGGAACCTCTCGGCAGCGGCCCGCGAGCGCGGT 2907
 QY 95 ----- 95
 Db 2908 CCGGTCGCGCGCGCGAGTAGTGTGCGCTTAGAGAGCTAGCGCGCCCGCGCG 2967
 QY 95 ----- 95
 Db 2968 GCCTCAAGTCAGCCGACTTGTCCCTGGGCGGCCACCTCACCTTCTCTTTTCTGTCT 3027
 QY 96 -----GlyAspValPheIleLeuValPheSerLeuAspAsnArgAspSe 110
 Db 3028 CTCTGTGCCCCCTTAGGAGAGCTTTTCATCTGTGTTCAGTCTGGACAAACCGCGACTC 3087
 QY 110 rPheGluGluValGlnArgLeuArgGlnIleLeuAspThrLysSerCysLeuLysAs 130
 Db 3088 CTTGAGGAGGTGCGAGCGGCTCAAGCAGCAGATCTCTGACACCAAGTCTTGCTCAAGAA 3147
 QY 130 nLysThrLysGluAsnValAspValProLeuValIleCysGlyAsnLysGlyAspArgAs 150
 Db 3148 CAAACCAAGGAGAACGTGGACGTGCTGCTGTCGCGCAACAAGGTGACCGCA 3207
 QY 150 pPheTyArgGluValAspGlnArgGluIleGluGlnLeuValGlyAspAspProGlnAr 170
 Db 3208 CTTCTACCCGAGGTGGACACGCGGAGATGAGCAGCTGTGGCGACGACCCCGACG 3267
 QY 170 gCysAlaTyPheGluIleSerAlaLysLysAsnSerSerLeuAspGlnMetPheArgAl 190
 Db 3268 CTGGCCTACTTCGAGATCTCGGCCAAGAAGAACAGCAGCTGGACACAGATGTTCCGCGC 3327
 QY 190 aLeuPheAlaMetAlaLysLeuProSerGluMetSerProAspLeuHisArgLysValSe 210
 Db 3328 GCTCTTGGCCATGGCCAGCTGCCAGCGAGATGAGCCAGACCTGCGACCCCAAGGTTCTC 3387
 QY 210 rValGlnTyrcysAspValLeuHisLysLysAlaLeuArgAsnLysLysLeuLeuArgAl 230
 Db 3388 GGTGAGTACTCGAGCTGCTGCAAGAAGCGCTCGGAACAGAGCTGTGCGCGGC 3447
 QY 230 aGlySerGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 250
 Db 3448 CGGCAGCG 3507

QY 250 aatgagatgproservalhisseraspbleumettrylleargluysalaseralaglyse 270
 Db 3508 GGGCCGGCCGACGATGATCAAGAGGACCTCATGATCCGCGAAGGACGCGCGGACG 3567
 QY 270 rgtlnalalysaplysaglunargysvalleser 281
 Db 3568 CCAAGCCAAAGACAGGACGCTGCTCATGACG 3601
 RESULT 13
 AL603710/c
 LOCUS AL603710 179124 bp DNA linear ROD 05-APR-2002
 DEFINITION Mouse DNA sequence from clone RP23-247B13 on chromosome 11,
 complete sequence.
 ACCESSION AL603710
 VERSION AL603710.8 GI:20608514
 KEYWORDS HTG;
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE
 AUTHORS Almeida J.
 TITLE Direct Submission
 JOURNAL Submitted (04-APR-2002) Wellcome Trust Sanger Institute, Hinxton,
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
 humbry@sanger.ac.uk; clone requests: clonerequests@sanger.ac.uk
 On Apr 7, 2002 this sequence version replaced gi:18070899.
 During sequence assembly data is compared from overlapping clones.
 Where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the
 variation annotation may not be found in the sequence submission
 corresponding to the overlapping clone, as we submit sequences with
 only a small overlap as described above.
 This sequence was finished as follows unless otherwise noted: all
 regions were either double-stranded or sequenced with an alternate
 chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by at least
 one plasmid subclone or more than one M13 subclone; and the
 assembly was confirmed by restriction digest. The following
 abbreviations are used to associate primary accession numbers given
 in the feature table with their source databases: Em, EMBL; SW,
 SWISSPROT; Tr, TrEMBL; Wp, WormPeP; Information on the WormPeP
 database can be found at
 http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-247B13 is
 from the RP23-23 Mouse PAC Library
 constructed by the group of Pieter de Jong.
 For further details see http://www.chori.org/bacpac/home.htm
 VECTOR: pBACe3.6.
 FEATURES
 source location/Qualifiers
 1..179124
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /chromosome="11"
 /clone="RP23-247B13"
 /clone_id="RP23-23"
 BASE COUNT 46162 a 43228 c 43736 g 45998 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 2,486-117 Length: 179124
 Score: 1370.00 Matches: 274
 Percent Similarity: 83.43% Conservative: 3
 Best Local Similarity: 82.53% Mismatches: 52
 Query Match: 94.48% Indels: 2
 DB: 10 Gaps: 2
 US-09-709-103-1F1 (1-282) x AL603710 (1-179124)
 QY 1 MetLysLeuAlaMetLysLysMetCysProSerAspSerGluLeuSerIlePro 20
 Db 122425 ATGAACTGGCGCGATGATCAAGAGATGTGCCAAAGGACTGTAAGTGTATCCCG 122466

QY 21 AlAlysAnCysTyrArgMetValIleLeuGlySerSerLysValGlyThrAlaIle 40
 Db 122365 GCCAAGAACTGTACAGAGATGTATCTCTGCTCATTCMAAGTGGCAAGACGGCATT 122306
 QY 41 ValSerArgPheLeuThrGlyArgPheGluAspAlaTyrThrProThrIleGluAspPhe 60
 Db 122305 GTGTGGGCTTCTTCACGGGCGCTTTCAGAGATGCTTACCCCTTACATCGAGGACTTC 122246
 QY 61 HisArgLysPheTyrSerIleArgGlyGluValIleTyrGluLeuAspIleLeuAspThrSer 80
 Db 122245 CACCGAAGATTATCTCATGATCCCGGAGAGTTTACCACTTGACATCTCGACATCTC 122186
 QY 81 GlyAsnHisProPheProAlaMetArgArgLeuSerIleLeuThr 95
 Db 122185 GGCATCATCGTTTCCCGCATGGCGGCTCTCTATCTCATCAAGTGAAGTGGAGTTC 122126
 QY 95 95
 Db 122125 GACGGGAGTATGGGGACCGGAGTCTTGAGGGAGTGAATGGGCTGTGTGTCTTG 122066
 QY 95 95
 Db 122065 GGGCTTCCCTTCACCTGGGGGTCTGCAGCTTGTCTGCTGGCAGCTTCTCACCTTT 122006
 QY 96 96
 Db 122005 CCACCTGTTCCCTTATAGAGAGCGTTTCTATCTGTGTGAGCTTATAGCAACCGCGAC 121946
 QY 110 SerPheGluGluValGlnArgLeuArgGlnIleLeuAspThrLysSerCysLeuLys 129
 Db 121945 TCATTCMAAGAGGGCAAGAGCTCAACAGCAATCTTACACCAAGCTGTCTCAAG 121886
 QY 130 AsnLysThrLysGluAsnValAspValProLeuValIleCysGlyAsnLysGlyAspArg 149
 Db 121885 AACAAACCAAGAGATGAGACGTGCGCTGTATTTGCGTAAACAAGGGACCG 121826
 QY 150 AspPheTyrArgGluValAspGlnArgGluIleGluGlnValGlyAspAspProGln 169
 Db 121825 GACTTCACCGGGAGAGTACGACGCGGAGATTTAGACAGCTGTGTGTACGACCTCAG 121766
 QY 170 ArgCysAlaTyrPheGluIleSerAlaLysLysAsnSerSerLeuAspGlnMetPheArg 189
 Db 121765 CGTTGTCTTACTTCGAGATCTCAGACCAAGAAACAGCAGCTTGGACAGATGTTCCCT 121706
 QY 190 AlaLeuPheAlaMetAlaLysLeuProSerGluMetSerProAspLeuHisArgLysVal 209
 Db 121705 GCGCTCTTTCCTCAGGCAAGCTGCTTACGAGATGAGCCCTGACCTTGACCGCAAGGTA 121646
 QY 210 SerValGlnTyrCysAspValLeuHisLysLysAlaLeuArgAsnLysLysLeuLeuArg 229
 Db 121645 TCTGTGCACTACTCGACAGCTACTGACACAGAAAGCTCTGAGAAACAAGAGCTTCTGCGT 121586
 QY 230 AlaGlySerGlyGlyGlyGlyGlyAspProGlyAspAlaPheGlyIleValAlaProPhe 249
 Db 121585 GCGGGCAGC---GGAGGCGGGGGGACCAAGCGGAGTCTTGTGCACTTGGCGCCCTTT 121529
 QY 250 AlaArgArgProSerValHisSerAspLeuMetTyrIleArgGluLysAlaSerAlaGly 269
 Db 121528 GCTTCGACAGCCAGCGGCGACACGAGCTTATGTATCTGTGTAACCAAGCAAGTGTGCG 121469
 QY 270 SerGlnAlaLysAspLysGluArgCysValIleSer 281
 Db 121468 AGCCAGGCTTAAGCAAGAGCGCTGTGTATCACT 121433
 RESULT 14
 AC025909/c
 LOCUS AC025909 247899 bp DNA linear HTG 27-JUN-2001
 DEFINITION Mus musculus chromosome 11 clone RP23-82E8, WORKING DRAFT SEQUENCE,
 6 unordered pieces.
 ACCESSION AC025909
 VERSION AC025909.28 GI:14547761
 KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.


```

|||||
Db 19030 GCGGCGAGC---GGAGCGCGGGGCGACACGCGGATGCTTGGATCTTGCGCCCTT 18974
QY 250 AAGATATGProserValHISserAPleMetTylIeAGTulValAseerAlaGly 269
Db 18973 GCTCGAGACCCGCGCTGACACGCGACCTCATGTATCATTCGTGMAAAACACAGTGTCCGC 18914
QY 270 SerGAlaIalysAspIysGluArgCysValIleSer 281
Db 18913 AGCCAGGCTTAGACAAAGAGCCCTGTCTCATCACT 18878

RESULT 15
AC090608 183598 bp DNA linear HTG 22-MAR-2002
LOCUS Homo sapiens chromosome 17 clone RP11-524F11 map 17, *** SEQUENCING
DEFINITION IN PROGRESS ***, 2 unordered pieces.
AC090608 AC090608.7 GI:19683171
VERSION HTG: HTGS PHASE1; HTGS_FULLTOP; HTGS_ACTIVEFIN.
KEYWORDS Homo sapiens.
SOURCE Homo sapiens.
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 183598)
AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.
TITLE Homo sapiens chromosome 17, clone RP11-524F11
JOURNAL Unpublished
2 (bases 1 to 183598)
REFERENCE 2 (bases 1 to 183598)
AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, S.,
Barna, N., Bastien, V., Boguslavsky, L., Boukhalter, B., Brown, A.,
Camaratia, J., Campopiano, A., Choepel, Y., Colangelo, M., Collins, S.,
Collymore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S.,
Dodge, S., Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J.,
Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Heatord, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,
Jones, C., Karatas, A., Lacroque, K., Lamazares, R., Landers, T.,
Lehoczyk, J., Levine, R., Liu, G., Maclean, C., Macdonald, P.,
Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K.,
McPheeters, R., Meidrim, J., Meneus, L., Mihova, T., Mlenga, V.,
Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C.H.,
O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,
Phunhthang, P., Pierre, N., Pollara, V., Raymond, C., Relta, R.,
Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetli, M.,
Roy, A., Santos, R., Schauer, S., Schupack, R., Seaman, S., Severy, P.,
Sougez, C., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
Travers, M., Travis, N., Triglio, J., Vassiliev, H., Viel, R., Vo, A.,
Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J.,
Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (03-MAR-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 183598)
REFERENCE 3 (bases 1 to 183598)
AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L.,
Boukhalter, B., Brown, A., Camaratia, J., Campopiano, A., Chang, J.,
Choazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A.,
Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., Lacroque, K., Lamazares, R.,
Landers, T., Lehoczyk, J., Levine, R., Lindblad-Toh, K., Liu, G.,
Maclean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,
McEwan, P., McEwan, P., McKernan, K., Meidrim, J., Meneus, L.,
Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R.,
Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,
Oliver, J., Peterson, K., Phunhthang, P., Pierre, N., Pollara, V.,
Raymond, C., Relta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
Roman, J., Rosetli, M., Roy, A., Santos, R., Schauer, S., Schupack, R.,
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Triglio, J., Vassiliev, H.,

```

```

Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (22-MAR-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 22, 2002 this sequence version replaced gi:16445176.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: MIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L11985
Center clone name: 524_F_11
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 178215: contig of 178215 bp in length
* 178216 178315: gap of 100 bp
* 178316 183598: contig of 5283 bp in length.
Location/Qualifiers
1. 183598
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/map="17"
/clone="RP11-524F11"
/clone_id="RP11-524F11"
Human Male BAC"
BASE COUNT 40950 a 47626 c 49185 g 45565 t 272 others
ORIGIN
Alignment Scores:
Pred. No.: 5.34e-116 Length: 183598
Score: 1356.00 Matches: 279
Percent Similarity: 78.81% Conservative: 0
Best Local Similarity: 78.81% Mismatches: 2
Query Match: 93.52% Indels: 73
Gaps: 1
Db: 1
US-09-709-103-1f1 (1-282) x AC090608 (1-183598)
QY 1 MetLysLeuAlaAlaMetIleLysLysMetCysProSerAspSerGluLeuSerIlePro 20
Db 181508 ATGAAATCTGCGCGCATATCAAGAAAGATGTGCCGAGCGACTGGAGCTGATATCCCG 181567
QY 21 AlAlysAsnCySTyTArMetValIleLeuGlySerSerLysValGlyVtThAlaIle 40
Db 181568 GCCAAGAACTGCTATGCGATGCTATCTCGCTCTGCTCCAGGTGGGCGAAGACCGCATC 181627
QY 41 ValSerArgPheLeuThyGlyArgPheGluAspAlaTyrrTrpProThrIleGluAspPhe 60
Db 181628 GTGTGCGGTTCTTCAACGGCGGCTTCAGAGAGCGCTTACAGCGCTTACATCAGAGACTTC 181687
QY 61 HisArgLysPheTySerIleArgGlyGluValTyrgInLeuAspIleLeuAspThsSer 80
Db 181688 CACCGCAATTTCATTCATCGCGGCGAGGCTTACAGCGCTTACATCCTCGACACGCTCC 181747
QY 81 GlyAsnHisProPheProAlaMetArgArgLeuSerIleLeuThr----- 95
Db 181748 GCGAACAACCCCTTCCCGCCGATGCGGCGCTTCCATCTCACAAGGTGAGCGGGGCC 181807
QY 95 ----- 95
Db 181808 GGGCAGGTGCGGAGAGGAGGCGGGAACCTCGGCCAGGCGCCCGCGAGCGCGGT 181867

```

QY 95 ----- 95
Db 181868 CCGGCTGCCGCGCGAGTAGTGGCTTCGGCGCTTAGAGAGGCTAGCGCGCCGCGCG 181927
QY 95 ----- 95
Db 181928 GCCTCAAGTCAGCCGACTTGTCCCTGGGGGCGCACCTCACCCTTCTCTTTCTGCT 181987
QY 96 -----GlyAspValPheIleLeuValPheSerLeuAspAsnArgAspSe 110
Db 181988 CTCTGTGCCCTCTAGGAGAGCTTTTCATCCTGTGTGTTCAGTCTGGACAACCGGACTC 182047
QY 110 rPheGluGluValGlnArgGlnGlnIleLeuAspThrLysSerCysLeuLysAs 130
Db 182048 CTTGAGGAGGTGCGACGGCTCAGCGCAGATCTCGACACCAAGTCTTGCTCAAGAA 182107
QY 130 nLysThrLysGlu--AsnValAspValProLeuValIleCysGlyAsnLysGlyAspArg 149
Db 182108 CAATACCAAGGAGANACGTNGACGTGCCCTGGTCATCTGCGGCAACAGGGGTGACCGC 182167
QY 150 AspPheTyrArgGluValAspGlnArgGluIleGluGlnLeuValGlyAspAspProGln 169
Db 182168 GACTTCTACCGCGAGGTGACCGCGAGATCGAGCAGCTGGTGGGCGACGCCCCAG 182227
QY 170 ArgCysAlaTyrPheGluIleSerAlaLysLysAsnSerSerLeuAspGlnMetPheArg 189
Db 182228 CGCTGCGCCTACTTCGAGATCTCGGCAAGAGAAACAGCAGCCTGGACCGATGTTCGC 182287
QY 190 AlaLeuPheAlaMetAlaLysLeuProSerGluMetSerProAspLeuHisArgLysVal 209
Db 182288 GCGCTCTTCGCCATGGCCAGCTGCCAGCGAGATGAGCCAGACCTGCACCGCAAGGTC 182347
QY 210 SerValGlnTyrCysAspValLeuHisLysLysAlaLeuArgAsnLysLysLeuLeuArg 229
Db 182348 TCGGTCCAGTACTGCGACGCTGTGCACAAGAAGCGCTGCGGAACAAGAGCTGCTGCGG 182407
QY 230 AlaGlySerGlyGlyGlyAspProGlyAspAlaPheGlyIleValAlaProphe 249
Db 182408 GCCGCTT 182467
QY 250 AlaArgProSerValHisSerAspLeuMetTyrIleArgGluLysAlaSerAlaGly 269
Db 182468 GCG 182527
QY 270 SerGlnAlaLysAspLysGluArgCysValIleSer 281
Db 182528 AGCCAGGCCAAGGACAGGAGCGCTGCGTCATCAGC 182563

Search completed: December 30, 2002, 15:42:49

Job time : 3326 secs